

**Konzeption zum optimierten Einsatz von Korrespondenzanalysen zur Position- und Lagebestimmung mit Hilfe von Kamerasystemen**

# **Masterthesis**

**zur Erlangung des akademischen Grades  
Master of Engineering  
an der  
Fachhochschule für Technik und Wirtschaft Berlin (FHTW)**

**Fachbereich Ingenieurwissenschaften I  
Studiengang Informationstechnik / Vernetzte Systeme**

**Hochschulbetreuer:  
Prof. Uwe Metzler  
Betrieblicher Betreuer:  
Dipl.-Ing. Denis Grießbach**

**vorgelegt von  
Eilke Santjer**

**erstellt beim Deutschen Zentrum für Luft- und Raumfahrt e.V. (DLR)**

**Berlin, 13.Mai.2008**

Begonnen: 8. Oktober 2007

Beendet: 13. Mai 2008

## Vorwort

Diese Arbeit behandelt den optimierten Einsatz von Algorithmen zur Merkmalsextraktion und Korrespondenzanalyse bei der Stereoanalyse von Bilddaten. Ziel dabei ist es aufgrund von Bewegungs- und Kameramessdaten eines möglichen Multisensoransatzes, korrespondierende Punktmengen für eine Zustandsschätzung mit hoher Effizienz und Qualität zu ermitteln. Es wird gezeigt wie Suchräume für mögliche Korrespondenzen in Bildfolgen mit Hilfe der Epipolargeometrie eingeschränkt werden können und somit der Arbeitsaufwand reduziert wird. Durch Integration von Bewegungsdaten können diese Verfahren nicht nur auf den Kameraaufbau, sondern auch auf unterschiedliche Aufnahmezeitpunkte angewendet werden. Es wird aufgezeigt welche Korrekturen bei Einsatz dieser Techniken notwendig sind und wie mit dem Brownsches-Verzeichnismodell der Einsatz realer Bilddaten ermöglicht wird. Zusätzlich Optimierungsansätze werden durch die Wiederverwendung von Merkmalen und durch den Einsatz einer Auflösungspyramide untersucht. Alle Optimierungsansätze wurden im praktischen Einsatz mit verschiedenen Operatoren getestet und die Auswirkungen auf Rechenaufwand, Zeiterparnis und Qualität dokumentiert.

Diese Arbeit entstand in Zusammenarbeit mit der Einrichtung „Optische Informationssysteme“ am Institut für Robotik und Mechatronik beim Deutschen Zentrum für Luft- und Raumfahrt e.V. in Berlin-Adlershof. Ich möchte mich dabei ganz besonders bei Denis Griebach für die großartige Unterstützung und lückenlose Betreuung bei dieser Masterarbeit bedanken und wünsche ihm alles Gute für seine Dissertation. Zusätzlich danke ich mich bei Dr. Anko Börner für seine Unterstützung und den ausgezeichneten Bedingungen zur Fertigung dieser Arbeit.

# Inhaltsverzeichnis

Abbildungsverzeichnis .....	VI
Tabellenverzeichnis .....	VII
Abkürzungsverzeichnis .....	VII
Nomenklatur .....	VIII
<b>1 Einführung .....</b>	<b>1</b>
<b>1.1 Motivation .....</b>	<b>1</b>
<b>1.2 Aufgabenstellung .....</b>	<b>2</b>
<b>1.3 Randbedingungen und Restriktionen .....</b>	<b>3</b>
<b>2 Grundlagen .....</b>	<b>4</b>
<b>2.1 Die Projektive Geometrie .....</b>	<b>4</b>
2.1.1 Projektiver Raum .....	4
2.1.2 Homogene Koordinaten .....	5
2.1.3 Direkte Lineare Transformation .....	6
2.1.4 Der Affine Raum .....	6
2.1.5 Euklidischer Raum .....	6
<b>2.2 Das Kameramodell .....</b>	<b>8</b>
2.2.1 Der Abbildungsprozess .....	9
2.2.2 Äußere Orientierung .....	10
2.2.3 Innere Orientierung .....	10
2.2.4 Verzeichnungen der Bildebene .....	12
<b>2.3 Stereoanalyse .....</b>	<b>13</b>
2.3.1 Stereogeometrie .....	13
2.3.2 Epipolar geometrie .....	14
2.3.3 Korrespondenzanalyse .....	16
<b>3 Analyse .....</b>	<b>17</b>
<b>3.1 Bestehende Lösung .....</b>	<b>17</b>
<b>3.2 Programmfunktionalitäten .....</b>	<b>18</b>
3.2.1 Initialisierung .....	18
3.2.2 Laden und Glättung der Bilddaten .....	18
3.2.3 Merkmalsextraktion .....	18

3.2.4	Korrespondenzanalyse durch Intra/Inter-Matching .....	19
3.2.5	Die Auflösungspyramide .....	21
3.2.6	Sonstige Funktionen.....	22
3.2.7	Allgemeiner Programmdurchlauf.....	22
3.3	<b>Problemanalyse.....</b>	<b>25</b>
4	<b>Systementwurf.....</b>	<b>26</b>
4.1	<b>Allgemeiner Lösungsansatz .....</b>	<b>26</b>
4.2	<b>Systemaufbau.....</b>	<b>27</b>
4.2.1	Bezugssysteme (Ttcp,Tcam,Tmov) .....	28
4.3	<b>Optimierung der Suchbereiche bei Matchingverfahren.....</b>	<b>31</b>
4.3.1	Berechnung der Epipolarlinie.....	31
4.3.2	Einschränkung des Tiefenbereichs.....	35
4.3.3	Nutzung der Ansätze der Epipolargeometrie .....	37
4.4	<b>Weiternutzung gefundener Merkmale .....</b>	<b>39</b>
4.4.1	Nutzung der Merkmalswiederverwendung .....	39
4.5	<b>Alternativer Matchingansatz durch 2. Inter-Matching.....</b>	<b>40</b>
4.6	<b>Verzeichnungskorrektur bei realen Bilddaten .....</b>	<b>41</b>
4.7	<b>Einflüsse auf den Pyramidenansatz.....</b>	<b>43</b>
5	<b>Testsoftware.....</b>	<b>44</b>
5.1	<b>Systemschnittstellen .....</b>	<b>44</b>
5.2	<b>Erweiterung in der Klassenhierarchie .....</b>	<b>46</b>
5.2.1	Die Kamera-Klasse .....	46
5.2.2	Die Klasse Epipolargeometrie.....	46
5.3	<b>Anpassungen in der algorithmischen Struktur .....</b>	<b>49</b>
5.4	<b>Funktionalitäten .....</b>	<b>50</b>
6	<b>Testanalysen und Verifikation.....</b>	<b>51</b>
6.1	<b>Versuchsbeschreibung .....</b>	<b>51</b>
6.1.1	Testsoftware .....	51
6.1.2	Testsystem.....	51
6.1.3	Testreihen.....	52
6.1.4	Durchführung und Auswertung.....	54

6.2	Gegenüberstellung des Epipolaransatzes zum Box-Verfahren.....	55
6.3	Einfluss der Optimierungsansätze .....	58
6.4	Einfluss der Optimierungsansätze mit andere Operatoren .....	63
6.5	Prüfung mit realen Bilddaten.....	66
7	Zusammenfassung.....	69
7.1	Erreichter Stand .....	69
7.2	Ausblick, Fortführung .....	71
	Literaturverzeichnis.....	72
	Erklärung zur Urheberschaft .....	73
	Anhang .....	74

## Abbildungsverzeichnis

Abbildung 1: Realisierungsansatz des „Integral Positioning System“ (IPS).....	2
Abbildung 2: Projektion einer Bildszene auf das Projektionszentrum .....	4
Abbildung 3: Modell einer Lochkamera .....	8
Abbildung 4: Darstellung des Mattscheibenmodells .....	9
Abbildung 5: Abbildungsvorgang.....	9
Abbildung 6: Externe Transformation (WKS in KKS) .....	10
Abbildung 7: verzerrtes Bild (links) und unverzerrtes Bild (rechts) .....	12
Abbildung 8: Darstellung der Epipolargeometrie .....	14
Abbildung 9: Zusammenhang zwischen Sichtstrahl und Epipolarlinie .....	15
Abbildung 10: Korrespondenzanalyse über 2 Bildpaare .....	17
Abbildung 11: Beispiel für ein 5x5 Pattern um das Pixel x:34 und y:67.....	19
Abbildung 12: Matchingalgorithmus der vorliegenden Testsoftware .....	20
Abbildung 13: Klassifizierung des Intra-/Inter-Matching .....	21
Abbildung 14: Beispiel für eine 3-stufige Auflösungspyramide zwischen 2 Bildern ...	22
Abbildung 15: Programmablauf der Originalsoftware .....	24
Abbildung 16: Systemmodell aus Stereokamerasystem und Bewegungssensorik .....	27
Abbildung 17: Zusammenhang zwischen den einzelnen Bezugssystemen.....	28
Abbildung 18: Abbildungsfehler bei Rasterung von Geraden (rot).....	33
Abbildung 19: Annäherung von Pixelgeraden nach Bresenham .....	34
Abbildung 20: Erweitertes Matching durch Nutzung der Epipolargeometrie .....	38
Abbildung 21: Wiederverwendung von Merkmalen .....	40
Abbildung 22: Matchingalternative durch zweites Inter-Matching .....	41
Abbildung 23: Kontext der Testsoftware .....	45
Abbildung 24: Implementierung der Klasse Kamera.....	46
Abbildung 25: Implementierung der Klasse Epipolargeometrie .....	48
Abbildung 26: Synthetische Bildszene aus der achsparallelen Versuchsreihe .....	52
Abbildung 27: Synthetische Bildszene aus dem realen Messdatensatz .....	53
Abbildung 28: Arbeitseinsparung der untersuchten Optimierungsschritte in % .....	65
Abbildung 29: Fehlerverhalten der untersuchten Optimierungsschritte in %.....	66

## Tabellenverzeichnis

Tabelle 1: Parameterbeschreibung der Klasse Epipolargeometrie.....	47
Tabelle 2: Kameraeigenschaften der synthetischen Bildreihen .....	52
Tabelle 3: Kameraeigenschaften der realen Messreihe.....	54
Tabelle 4: Parametrierung – Versuch I .....	55
Tabelle 5: Suchfenstergrößen für Intra/Inter-Matching - Versuch I .....	56
Tabelle 6: Matching-Ergebnisse zwischen Box-Ansatz und Epipolarlösung .....	56
Tabelle 7: Allgemeine Systemparameter .....	58
Tabelle 8: Abweichende Parametrierungen für den Epipolarsuchbereich .....	59
Tabelle 9: Ergebnisse: gesamte und eingeschränkte Epipolarlinie .....	60
Tabelle 10: Ergebnisse: Merkmalswiederverwendung und 2. Inter-Matching .....	61
Tabelle 11: Parametrierung bei Gegenüberstellung der Operatoren.....	63
Tabelle 12: Matching-Ergebnisse bei verschiedenen Operatoren.....	64
Tabelle 13: Parametrierung bei Untersuchung der realen Bildfolge.....	67
Tabelle 14: Matchingergebnisse bei realer Bildfolge .....	67

## Abkürzungsverzeichnis

WKS	Weltkoordinatensystem
KKS	Kamerakoordinatensystem
IMU	Inertial Measurement Unit
TCP	Tool Centre Point
KLT	Kanade-Lucas-Tomasi
NCC	Normalized cross correlation
SAD	sum of absolute differences
SSD	sum of squared differences
IPS	Integral Positioning System
GPS	Global Positioning System
CCD	Charge-coupled Device

## Nomenklatur

$x, y, z$	Koordinaten eines 3D-Punkts (Indizierungen: c Kamerakoordinaten, w Weltkoordinaten, o Objektkoordinaten)
$\omega, \varphi, \kappa$	Rotationswinkel
$X, Y, Z, W$	Objektpunkt in homogenen Koordinaten
$t_n$	Zeitpunkt
$p$	2D-Vektor
$P$	3D-Vektor
$\sim$	Vektor in homogenen Koordinaten
$A$	Matrix
$A^{-1}$	Inverse Matrix von A
$A^T$	Transponierte Matrix
$H$	Homographie-Matrix
$R$	Rotationsmatrix
$t$	Translationsvektor
$K$	intrinsische Transformationsmatrix/Kameramatrix
$u, v$	Pixelkoordinaten, Bildkoordinaten
$k_u, k_v$	horizontaler und vertikaler Skalierungsfaktor
$f$	Brennweite
$C$	Fokales Zentrum der Kamera, Brennpunkt
$F$	Fundamentalmatrix
$l$	Epipolarlinie
$e$	Epipol
$[.]_x$	asymmetrische Matrix eines Vektors
$dT$	4x4 euklidische 3D-Transformationsmatrix
$m$	Geradensteigung
$d_u, d_v$	Differenz-Bereich in Pixelkoordinaten
$d_{mi}, d_{max}$	Mindest- und Maximalabstand von Objektpunkten
$\delta_u, \delta_v$	Disparität, Verzerrung
$K_n$	Verzeichnungsparameter
$T_{cam}$	relative Orientierung zwischen den Kameras



# 1 Einführung

Algorithmen zur visuellen Wahrnehmung des Umfelds, finden heute in verschiedenen Bereichen der Industrie und Forschung ein hohes Interesse. Die Thematik, mit Hilfe des Computers, aus natürlichen Ansichten eines Kamerasystems, physikalische Merkmale zu Position und Lage zu bestimmen, hat eine hohe Aktualität. Als Beispiele seien hier die Entwicklung autonomer Fahrsysteme oder Forschungen an visuellen Orientierungshilfen von Robotern erwähnt. Wie wäre es z.B., wenn ein Roboter sich aufgrund seiner optischen Sensoren innerhalb von Räumen ohne Vorkenntnisse orientieren könnte? Hierzu müsste er allerdings, die von ihm erfassten Bilder auf Bewegungseindrücke und räumliche Dimensionen beurteilen können. Im Rahmen eines Projekts beim Deutschen Zentrum für Luft- und Raumfahrt e.V. sollen entsprechende Ansätze näher betrachtet werden.

## 1.1 Motivation

Für viele Messaufgaben werden für eine zeitliche und/oder räumliche Referenzierung bis zu sieben physikalische Größen benötigt. Zu diesen gehören Position ( $x,y,z$ ), Lage ( $\omega,\varphi,\kappa$ ) und Zeit ( $t$ ). Einsatzgebiete finden sich dabei in verschiedenen Beispielen, wie bei

- Abbildenden Sensoren in der Photogrammetrie und Fernerkundung
- Robotern in den Bereichen Montage, Medizin und Service
- Autonomen Fahrzeugen und Flugsystemen

Die messtechnische Bestimmung solcher Größen basiert in der Regel auf Lösungen für dedizierte Aufgaben, wie z.B. Navigationssysteme in Kraftfahrzeugen. Meßsysteme dieser Art basieren dabei meist auf Einzelsensoren.

Das Projekt mit der Bezeichnung „**Integral Positioning System**“ (IPS) beschäftigt sich mit einem Multi-Sensor-Ansatz zur Bestimmung von Position und Lage. Der Aufbau eines entsprechenden Sensorpools ermöglicht die Nutzung verschiedener Sensor-Technologien und Methoden zur Datenfusion. Als Sensoren sollen Kameras, Inertialmeßsysteme und laufzeit- bzw. phasenmessende Mikrowellensysteme (GPS, Pseudoliten) verwendet werden. Das Kernthema des Projekts ist die Entwicklung komplexer Algorithmen zur Zustandsschätzung. Der Fokus liegt auf Applikationen im Indoor- und Outdoor-Bereich.

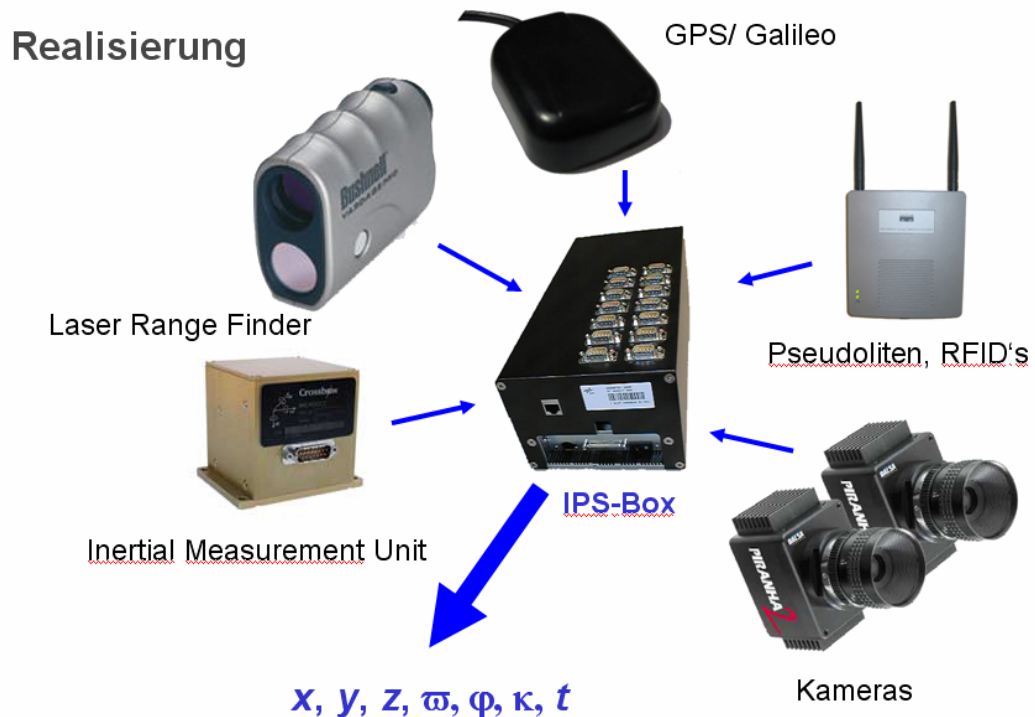


Abbildung 1: Realisierungsansatz des „Integral Positioning System“ (IPS)

Optische Sensoren spielen innerhalb des IPS-Konzeptes eine entscheidende Rolle. Aufgenommene Bildinformationen sollen genutzt werden, um Bewegungen der Kameras zu schätzen. Diese Informationen können dann mit Daten anderer Sensoren intelligent verknüpft werden.

## 1.2 Aufgabenstellung

Diese Arbeit verfolgt das Ziel den Einsatz von Algorithmen zur Merkmalsextraktion und Korrespondenzanalyse bei der Stereoanalyse von Bilddaten zu optimieren. Aufgabe solcher Verfahren ist es, eindeutige Merkmale in einem Bild zu extrahieren und über mehrere Bildsequenzen zu verfolgen. Dafür müssen korrespondierende Punktpaare in den Bildern gefunden werden. Aufgrund der Ergebnismengen können dann mit komplexen Algorithmen zur Zustandsschätzung Aussagen zu Position- und Lage des Kamerasystems gemacht werden.

Durch Kenntnis über relative Lage und Eigenschaften zweier Kameras können Ähnlichkeitsbedingungen zwischen den Ansichten mit Hilfe der Epipolargeometrie formuliert werden. Die Verwendung von Stereokameras gestattet die Ableitung von Tiefeninformationen aus Bilddaten. Dieser Umstand soll ausgenutzt werden die Suchbereiche für Korrespondenzanalysen zu bestimmen und optimal anzupassen. Verschiedene Optimierungsschritte werden dazu in ihren Ansätzen untersucht und

Auswirkungen in praktischen Versuchen nachgewiesen. Schnelligkeit, Robustheit und Qualität bei der Suche von Korrespondenzen stehen dabei im Vordergrund.

In einem ersten Schritt wird gezeigt, wie durch Ansätze der Epipolargeometrie, Suchräume zur Findung von Korrespondenzen bestimmt werden können. Die Betrachtung des Problems erfolgt zum einen bei Analysen zwischen den Kameraansichten zu einem Zeitpunkt und zum anderen bei unterschiedlichen Aufnahmezeitpunkten des Stereosystems. Dabei wird im zweiten Fall neben Lage und Eigenschaften der Kameras auch die Nutzung einer Bewegungsschätzung durch ein externes Sensorsystem berücksichtigt. Durch Einschränkung des Tiefenbereichs wird der Ansatz weiter verfeinert und optimiert.

Zusätzlich wird ein Verfahren zur Wiederverwendung von Merkmalen erarbeitet. Es wird aufgezeigt wie bereits gewonnene Vorkenntnisse aus vorangegangenen Analysen wiederverwendet und somit Einsparungen durch Verringerung des Rechenaufwands ermöglicht werden.

Durch Modifikation eines vorhandenen Testprogramms zur Korrespondenzanalyse werden die erarbeiteten Ansätze praktisch getestet und deren Verhalten untersucht. Durch Integration des Brownsches Verzeichnungsmodells [ZHA98] erfolgt der Nachweis neben synthetischen Bildfolgen auch mit realen Bilddaten.

### **1.3 Randbedingungen und Restriktionen**

Diese Masterarbeit baut auf Ergebnisse einer bereits beim DLR angefertigten Diplomarbeit mit dem Thema „Feature Detection und Matching Verfahren zur Positions- und Lagebestimmung“ [ROS06] auf. Diese Arbeit stellt bereits eine ausführliche Untersuchung unterschiedlicher Algorithmen zur Merkmalsextraktion und Korrespondenzanalyse bereit, so dass eine Klassifizierung und Analyse dieser Algorithmen nicht Bestandteil dieser Arbeit sind. Zusätzlich werden keine Verfahren zur Bewegungsschätzung oder zur Kamerakalibration behandelt. Die in dieser Arbeit diesbezüglich herangezogenen Daten beruhen auf praktischen Ergebnissen und können in dieser Form als real reproduzierbar angenommen werden. Trotz des praktischen Nachweises, ist der Einsatz der Testsoftware in dieser Form unter realen Bedingungen nicht vorgesehen.

## 2 Grundlagen

Zunächst sollen die theoretischen Grundlagen der Arbeit beleuchtet werden. Hierzu gehören relevante Grundlagen zur projektiven Geometrie, zum Kameramodell und zur Kalibrierung, sowie der Stereoanalyse mit Extraktions- und Matchingverfahren.

### 2.1 Die Projektive Geometrie

Grundlage der Projektiven Geometrie bildet die Perspektivische Projektion, welche sich mit der korrekten Darstellung einer dreidimensionalen Szene auf eine zweidimensionale Ebene befasst. Wird eine räumliche Szene von einem bestimmten Punkt aus betrachtet, konvergieren die Sichtstrahlen bis sie sich im Betrachtungspunkt schneiden.

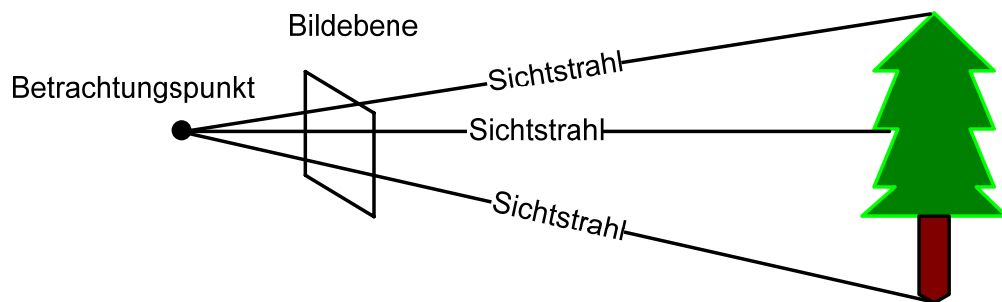


Abbildung 2: Projektion einer Bildszene auf das Projektionszentrum

Wird zusätzlich zwischen der darzustellenden Szene und dem Betrachtungspunkt senkrecht eine Bildebene eingefasst, ist ersichtlich dass alle Punkte eines Objekts sich auf der Bildebene perspektivisch richtig darstellen. Parallel verlaufende Linien im Raum konvergieren auf der Bildebene in sogenannten Fluchtpunkten. Größenverhältnisse werden auf der Bildebene dadurch nicht mehr korrekt abgebildet und die Darstellung gleichgroßer Objekte wird mit zunehmender Entfernung immer kleiner. Durch eine perspektivische Projektion ändern sich somit die Längen- und Flächenverhältnisse eines Objekts, wodurch innerhalb der Ebene ein räumlicher Eindruck entsteht.

#### 2.1.1 Projektiver Raum

Im Allgemeinen definiert sich ein geometrischer Raum durch seine Transformationsmöglichkeiten und den damit verbundenen Invarianten. Der projektive Raum beschreibt die allgemeinste Darstellung eines geometrischen Raums. In ihm sind die Transformationsmethoden, wie Rotation, Translation, Skalierung, Scherung und die mathematischen Gesetzmäßigkeiten der perspektivischen Projektion berücksichtigt.

Dieser Umstand verringert die Anzahl der unveränderlichen Eigenschaften, der sogenannten Invarianten, des Systems. Eine 4x4 Transformationsmatrix für den Projektiven Raum ist bis auf einen Skalierungsfaktor definiert und besitzt somit nur eine Invariante. Es verbleiben damit 15 Freiheitsgrade für eine Transformationsbeschreibung im projektiven Raum.

Ein entscheidender Vorteil der Darstellung im projektiven Raum liegt darin, Transformationen zwischen Koordinatensystemen zu vereinfachen und somit zusammenzufassen. Die projektive Transformation kann linear in homogenen Koordinaten beschrieben werden und ist invertierbar.

### 2.1.2 Homogene Koordinaten

Die Einführung von homogenen Koordinaten vereinfacht die Darstellung von räumlichen Beziehungen, durch deren Übertragung in den projektiven Raum. Der wesentlichste Vorteil ist es, dass auch Objekte im Unendlichen dargestellt werden können.

Um einen dreidimensionalen Punkt in homogenen Koordinaten anzugeben wird dieser um eine Dimension (T) erweitert. Durch diesen Schritt ist nur noch die Richtung und nicht mehr die Länge des Vektors relevant. Ein 3-dimensionaler Objektpunkt wird somit in homogenen Koordinaten mit 3+1 Komponenten abgebildet und ist bis auf einen Skalierungsfaktor vollständig beschrieben. Standardmäßig wird dabei diese zusätzliche Komponente zu 1 gesetzt. Alle Vektoren

$\lambda \cdot m \forall \lambda \neq 0$  beschreiben im projektiven Raum den gleichen Punkt. Ein Objektpunkt in homogenen Koordinaten wird mit einer Tilde ( $\sim$ ) dargestellt.

Euklidischer Objektpunkt in homogenen Koordinaten:

$$\tilde{m} = \lambda \cdot (x, y, z, 1)^T = (X, Y, Z, W)^T \quad (2-1)$$

Rücktransformation:

$$x = \frac{X}{W} \quad y = \frac{Y}{W} \quad z = \frac{Z}{W} \quad \text{für } W \neq 0$$

$$P = [x \quad y \quad z]^T \quad \rightarrow \quad \tilde{P} = [X \quad Y \quad Z \quad W]^T \quad (2-2)$$

### 2.1.3 Direkte Lineare Transformation

Ein 3-dimensionaler Szenepunkt in homogenen Koordinaten kann mathematisch durch eine direkte lineare Transformation (DLT) auf einen 2-dimensionalen Bildpunkt in homogenen Koordinaten abgebildet werden. Eine lineare Transformation hat dabei die allgemeine Form,

$$\tilde{p}_2 = H * \tilde{p}_1 \quad (2-3)$$

wobei  $H$  die Transformationsmatrix,  $\tilde{p}_1$  den Ausgangspunkt und  $\tilde{p}_2$  den Punkt nach der Transformation beschreibt. Zusätzlich kann das Produkt zweier linearer Transformationen wieder als lineare Transformation dargestellt werden. In einer DLT-Matrix können somit alle benötigten Informationen über Orientierung und Lage des Kamerasystems und der Abbildungseigenschaften zusammengefasst werden.

### 2.1.4 Der Affine Raum

Der affine Raum definiert sich als Unterraum des Projektiven Raums und stellt somit eine Einschränkung der Transformationsdarstellung zwischen zwei Koordinatensystemen dar. Die Umsetzung erfolgt durch eine lineare Transformation mit eingeschränkten Freiheitsgraden. Affine Transformationen ermöglichen neben Rotation und Translation, eine Skalierung oder Scherung, wobei Längen und Winkel aber nicht mehr übereinstimmen müssen. Parallelität und Längenverhältnisse bleiben allerdings erhalten und erweitern die Anzahl der Invarianten. Die affine Transformationsmatrix ist im Folgenden dargestellt:

$$T_A = \begin{bmatrix} A^n & b^n \\ 0_n^T & 1 \end{bmatrix} \quad (2-4)$$

### 2.1.5 Euklidischer Raum

Der euklidische Raum beschreibt die Geometrie unserer dreidimensionalen Welt. In ihm lassen sich Lage und Orientierung eindeutig beschreiben. Ein Punkt ist für drei Dimensionen definiert und wird durch einen 3-komponentigen Ortsvektor  $(x,y,z)^T$  dargestellt. Eine vollständige Transformation im euklidischen Raum kann durch folgenden Term beschrieben werden:

$$\vec{P}_2 = R_{3 \times 3} * \vec{P}_1 + \vec{t} \quad (2-5)$$

Der euklidische Raum definiert sich ebenfalls als Unterraum des projektiven Raums und kann somit durch homogene Koordinaten in diesen überführt werden. Die Freiheitsgrade einer Transformation bestehen dabei nur noch aus einer Rotation und Translation.

Die Translation beschreibt die Parallelverschiebung eines Punktes  $P_1$  durch einen Differenzvektor  $t$ . Der neu entstandene Punkt  $P_2$  ermittelt sich somit aus den Teilkomponenten der Ausgangslage und der Verschiebung selbst.

$$\vec{P}_2 = \vec{P}_1 + \vec{t} \quad (2-6)$$

Die Rotation beschreibt die Drehwinkel eines Vektors zu den drei Achsen des kartesischen Koordinatensystems. Die Drehung eines Punkts entsprechend eines Winkels  $(\omega, \varphi, \kappa)$  kann für jede Achse in einer orthogonalen  $3 \times 3$  Rotationsmatrix beschrieben werden.

$$R(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{bmatrix} \quad \text{Rotation um die x-Achse} \quad (2-7)$$

$$R(\varphi) = \begin{bmatrix} \cos \varphi & 0 & \sin \varphi \\ 0 & 1 & 0 \\ -\sin \varphi & 0 & \cos \varphi \end{bmatrix} \quad \text{Rotation um die y-Achse} \quad (2-8)$$

$$R(\kappa) = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{Rotation um die z-Achse} \quad (2-9)$$

Die resultierende Rotation ergibt sich unter Berücksichtigung der Reihenfolge aus der Multiplikation aller drei Teilrotationen.

$$R = R(\omega) \bullet R(\varphi) \bullet R(\kappa)$$

Die resultierende Rotation und die additive Translation kann durch eine gemeinsame lineare Transformationsmatrix in den Projektiven Raum überführt werden. Bei Berücksichtigung der Reihenfolge aus Rotationen und Translation ergibt sich, durch die nicht kommutative Matrizenmultiplikation, die lineare Transformationsmatrix in homogenen Koordinaten für den euklidischen Raum:

$$T_E = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \quad (2-10)$$

## 2.2 Das Kameramodell

Zur Abbildung einer räumlichen Szene auf eine zweidimensionale Bildebene kommt das Lochkameramodell zum Einsatz. Für die Bild- und Objektkoordinaten wird ein rechtshändiges kartesisches Koordinatensystem angenommen. Punkte einer dreidimensionalen (3D) Bildszene werden über den Brennpunkt (C), als zweidimensionale (2D) Punkte innerhalb einer Bildebene abgebildet. Die aufgenommene Abbildung wird dabei punktgespiegelt auf die dahinter liegende Bildebene projiziert.

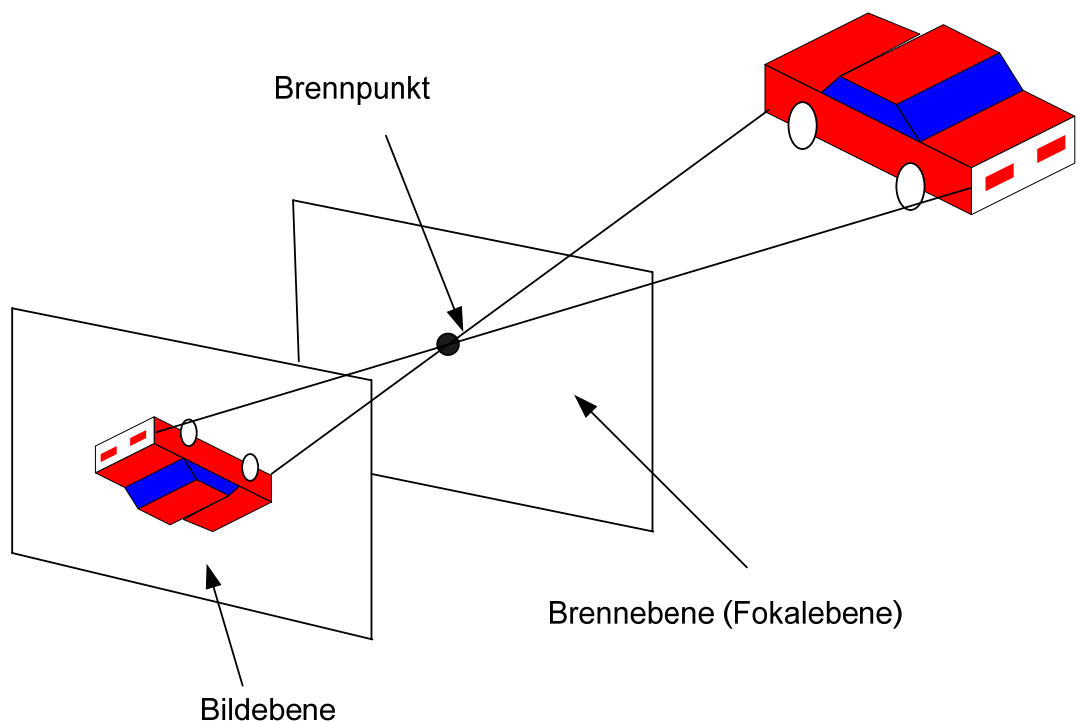


Abbildung 3: Modell einer Lochkamera



Da Objekte im Lochkameramodell auf der Bildebene spiegelverkehrt dargestellt werden und dieser Umstand die grafische und geometrische Darstellung unnötig verkompliziert, kommt das Mattscheibenmodell zum Einsatz. Dieses Modell unterscheidet sich dabei lediglich durch die Verschiebung des Projektionspunkt hinter die Bildebene, wobei allerdings x- und y-Achsen des Koordinatensystems gespiegelt werden.

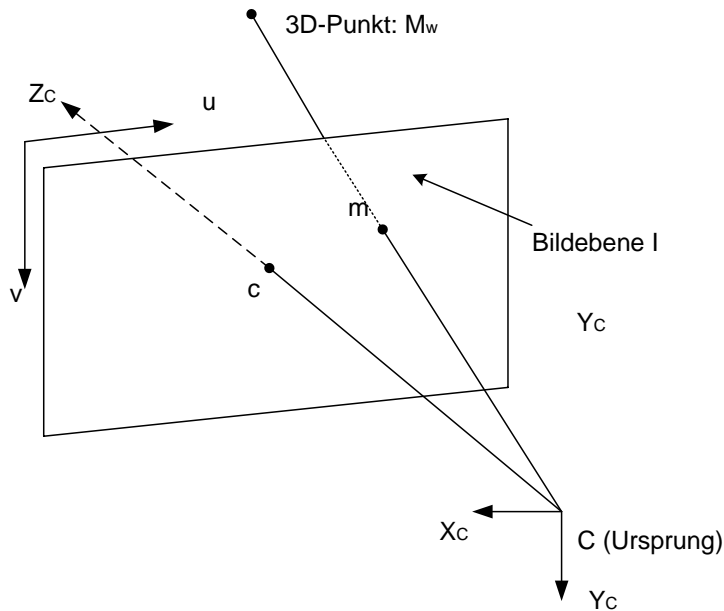


Abbildung 4: Darstellung des Mattscheibenmodells

### 2.2.1 Der Abbildungsprozess

Der Abbildungsprozess einer Kamera kann in zwei Stufen beschrieben werden. In der ersten Stufe wird das Kamerasystem mit dem Weltkoordinatensystem in Bezug gebracht. Die zweite Stufe beschreibt den Transformationsprozess der Szenenpunkte aus dem Kamerakoordinatensystem in die Bildkoordinaten. Die Darstellung des Abbildungsvorgangs einer Bildaufnahme beschreibt folgende Abbildung:

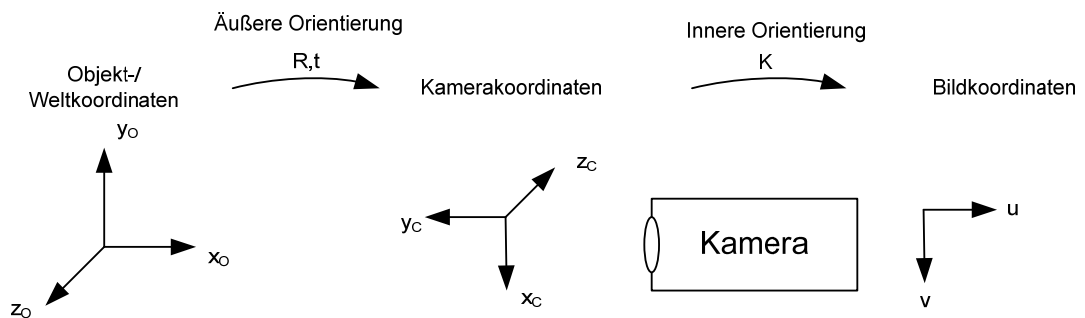


Abbildung 5: Abbildungsvorgang

### 2.2.2 Äußere Orientierung

Die äußere bzw. extrinsische Orientierung beschreibt die Positionierung und Ausrichtung des Kamerasystems im Objekt- bzw. Weltkoordinatensystem. Das Weltkoordinatensystem (WKS) bildet mit seinem Ursprung (O) das Bezugskoordinatensystem aller Objektpunkte einer Bildszene. Der Ursprung (C) des Kamerakoordinatensystems bildet das Projektionszentrum der Kamera. Um einen Objektpunkt auf die Bildebene zu projizieren, muss dieser in einem ersten Schritt in das Kamerakoordinatensystem transformiert werden. Eine euklidische Transformationsmatrix beschreibt die Umformung eines Objektpunkts des Weltkoordinatensystems in einen Punkt innerhalb der Kamerakoordinaten.

$$\tilde{P}_C = T_E \bullet \tilde{P}_O \quad (2-11)$$

Die Parameter dieser Transformation werden extrinsische Parameter der Kamera genannt.

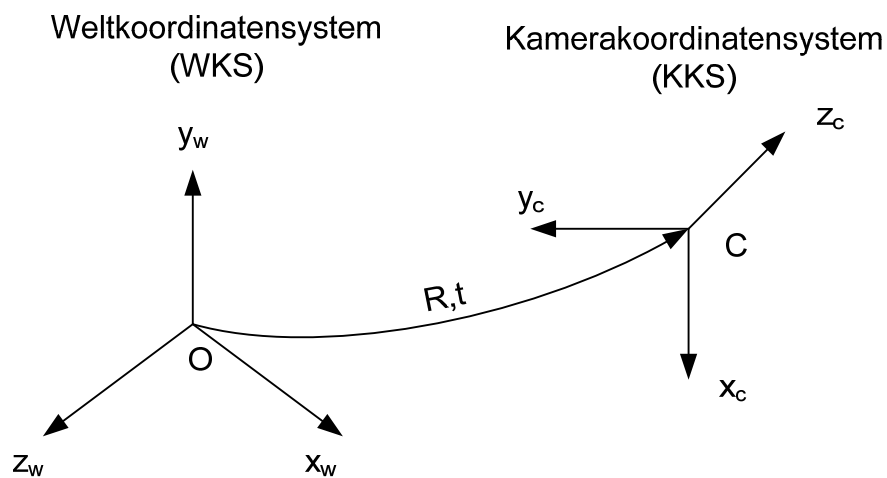


Abbildung 6: Externe Transformation (WKS in KKS)

### 2.2.3 Innere Orientierung

Ist eine Bildszene in das Kamerakoordinatensystem transformiert, so müssen die darin enthaltenen 3D-Punkte weiter in die 2D Bildebene überführt werden. Dieser Vorgang ist durch die perspektivische Projektion beschrieben. Die Grundlage bilden die Gleichungen der Zentralprojektion. Die Abbildungsvorschriften beschreiben die Transformation eines Objektpunkts des Kamerasystems  $P_c = (x_c, y_c, z_c)^T$  in einen Bild-

punkt  $(u, v)^T$  der Bildebene, wobei  $f$  den Abstand zwischen der Bildebene und dem Projektionszentrum entlang der optischen Achse beschreibt.

$$\frac{u}{x_c} = \frac{v}{y_c} = \frac{f}{z_c} \quad (2-12)$$

Die Eigenschaften einer realen Kamera definieren sich hauptsächlich über die Brennweite ( $f_{x,y}$ ), den Bildhauptpunkt ( $u_0, v_0$ ), und die Pixelscherung ( $\gamma$ ). Die Pixelscherung kann bei CCD Kameras, wegen der hohen Genauigkeit bei der Fertigung vernachlässigt werden und wird deshalb nicht weiter berücksichtigt. Da die Angabe der Brennweite meist metrisch verfolgt, ist noch die weitere Umrechnung in Bildkoordinaten notwendig. Entsprechende Skalierungsfaktoren ( $k_u, k_v$ ) ergeben sich aus der Pixelgröße des CCD Sensors. Die Verschiebung in den Koordinatenursprung des Bildes (meist in der linken oberen Ecke) wird über die Verschiebungsparameter  $u_k, v_k$  angegeben. Zusammenfassend kann die innere Orientierung des Kameramodells durch die intrinsische Transformationsmatrix  $K$  vollständig beschrieben werden.

$$K = \begin{bmatrix} fk_u & 0 & u_k \\ 0 & fk_v & v_k \\ 0 & 0 & 1 \end{bmatrix} \quad (2-13)$$

Die Parameter dieser Matrix können mit Hilfe eines Kalibrationsvorgangs ermittelt werden. Die dabei ermittelten Werte werden interne oder intrinsische Parameter der Kamera genannt. Die Transformation eines homogenen Objektpunkts von Kamerakoordinaten in Bildkoordinaten ist somit wie folgt beschrieben:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{4 \times 3} \bullet \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{mit} \quad K_{4 \times 3} = [K \mid 0] \quad (2-14)$$

## 2.2.4 Verzeichnungen der Bildebene

Bei einem realen Kamerasystem sind bei einem Abbildungsprozess zusätzlich nicht-lineare Verzerrungen zu berücksichtigen. Der entstandene Abbildungsfehler wird geometrisch bedingt durch die Bauweise der Kamera hervorgerufen. Großen Einfluss nimmt dabei die Krümmung der Linse, wodurch hauptsächlich in radialer Richtung Verzerrungen hervorgerufen werden. Das hat zur Folge, dass mit zunehmendem radialem Abstand zum Brennpunkt Bildpunkte innerhalb eines nichtlinear ansteigenden Fehlers verschoben abgebildet werden. Der Fehler äußert sich in der gesamten Abbildung durch eine Bildkrümmung, so dass z.B. gerade Kanten in den Außenbereichen gekrümmt dargestellt werden.



Abbildung 7: verzerrtes Bild (links) und unverzerrtes Bild (rechts)

Der Fehler kann im Allgemeinen durch eine Transformation von verzerrten in unverzerrte Bildkoordinaten beschrieben werden.

$$u = u_d + \delta \quad (2-15)$$

$$v = v_d + \delta \quad (2-16)$$

Die Koordinaten  $u_d$  und  $v_d$  beschreiben die verzeichneten (engl. distort) und  $u$  und  $v$  die unverzeichneten Bildkoordinaten in entsprechende Koordinatenrichtung. Die Verzerrung bzw. Verzeichnung wird üblicherweise als symmetrisch angesehen [Zha98] und ist durch den Verzeichnungsparameter  $\delta$  beschrieben.  $\delta$  besteht im Normalfall neben radialen auch aus tangentialen Verzeichnungskomponenten der Bildpunkte. Allerdings ist der tangentielle Anteil dabei so gering, dass er als Fehler keinen bedeutenden Einfluss nimmt und somit vernachlässigt werden kann. Die berücksich-

tigte radiale Verzeichnung kann mit Hilfe zu bestimmender Koeffizienten, sogenannter K-Parameter, über einen nichtlinearen Term beschrieben werden.

$$\delta_u = u_d(K_1 r_d^2 + K_2 r_d^4 + \dots) \quad (2-17)$$

$$\delta_v = v_d(K_1 r_d^2 + K_2 r_d^4 + \dots) \quad (2-18)$$

$$\text{mit } r_d = \sqrt{u_d^2 + v_d^2} \quad (2-19)$$

Durch Angabe der ersten drei K-Parameter wird der radiale Verzeichnungsfehler ausreichend berücksichtigt und ermöglicht somit die Korrektur von verzeichneten Bildpunkten in ein unverzeichnetes Bild. Die Parameter selbst werden bei der Kalibration der Kamera bestimmt.

## 2.3 Stereoanalyse

Die Stereoanalyse befasst sich im Allgemeinen mit der Suche von Korrespondenzen einer Bildszene aus zwei Perspektiven. Der Begriff Perspektive kann dabei sowohl als Sicht auf eine Szene zu unterschiedlichen Zeitpunkten, als auch als Sicht von unterschiedlichen Orten zum gleichen Zeitpunkt verstanden werden. Die Korrespondenzanalyse ermöglicht es, den Zusammenhang gemeinsamer Bildmerkmale bzw. Bildpunkte in beiden Perspektiven über Bewegungsvektoren darzustellen. Aus der Stereoanalyse können somit allgemeine Rückschlüsse über die Bewegung gezogen werden.

### 2.3.1 Stereogeometrie

Die Thematik der Stereogeometrie befasst sich mit der Anordnung und den damit verbundenen Beziehungen zwischen zwei Kameras. Solche Systemanordnungen werden Stereokamerasysteme genannt. Eine relative Orientierung zwischen beiden Kameras kann über eine Rotation  $R$  und eine Translation  $t$  beschrieben werden, wobei der Ursprung eines Kamerasystems als Bezugssystem angenommen wird. Die Transformationsgleichung des optischen Zentrums der ersten Kamera  $C_1$  in das optische Zentrum der zweiten Kamera  $C_2$  ergibt sich wie folgt:

$$\vec{C}_2 = R_{3 \times 3} * \vec{C}_1 + t \quad (2-20)$$

Ist die relative Lage und Richtung der beiden Bildkoordinatensysteme zueinander bekannt, so lässt sich der Suchbereich für einen extrahierten Bildpunkt im zweiten Bild erheblich einschränken.

### 2.3.2 Epipolargeometrie

Die Epipolargeometrie beschreibt die Zusammenhänge, die sich aus Orientierung und Lage zweier Kameras ergeben. Die Verbindung zwischen den optischen Zentren wird Basislinie  $B$  genannt. Schneidet die Basislinie, aufgrund eines Rotationswinkels zwischen beiden Kameras, beide Bildebenen, so werden die entstandenen Schnittpunkte Epipole genannt. Die Lage der Epipole in der Bildebene wird dabei nur von der Orientierung der Kameras zueinander bestimmt.

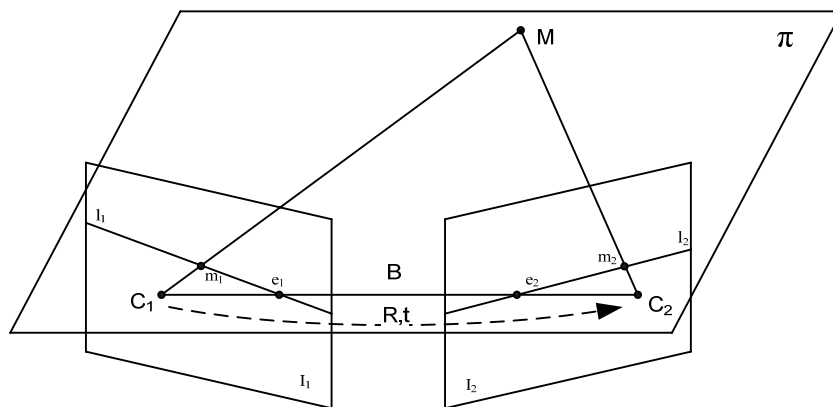


Abbildung 8: Darstellung der Epipolargeometrie

Der abzubildende Objektpunkt  $M$  spannt mit den beiden optischen Zentren  $C_1$  und  $C_2$  die Epipolarebene  $\pi$  auf. Die Linien  $l_1$  und  $l_2$  bilden dabei die Schnittgeraden zwischen der Epipolarebene und den beiden Bildebenen. Die Punkte  $m_1$  und  $m_2$  ergeben die Abbildung des Punktes  $M$  auf den jeweiligen Bildebenen.

Die Epipole beschreiben in ihrer zugehörigen Bildebene die Brennpunkte der jeweils anderen Kamera. Der Epipol  $e_2$  bildet den Brennpunkt  $C_1$  in Ebene  $I_2$  ab und umgekehrt  $e_1$  den Brennpunkt  $C_2$  in Ebene  $I_1$ .

Betrachtet man nun den Sichtstrahl zwischen dem Brennpunkt  $C_1$  und dem Objektpunkt  $M$ , so schneidet dieser die Bildebene  $I_1$  im Punkt  $m_1$ . Wandert der Punkt  $M$  auf dem Sichtstrahl in Richtung Kamera 1, so bleibt die Abbildung  $m_1$  in der Bildebene  $I_1$  erhalten. Auf der Ebene  $I_2$  ist jedoch zu beobachten, dass der Abbildungspunkt  $m_2$  sich entlang der Epipolarlinie  $l_2$  in Richtung des Epipols  $e_2$  bewegt.

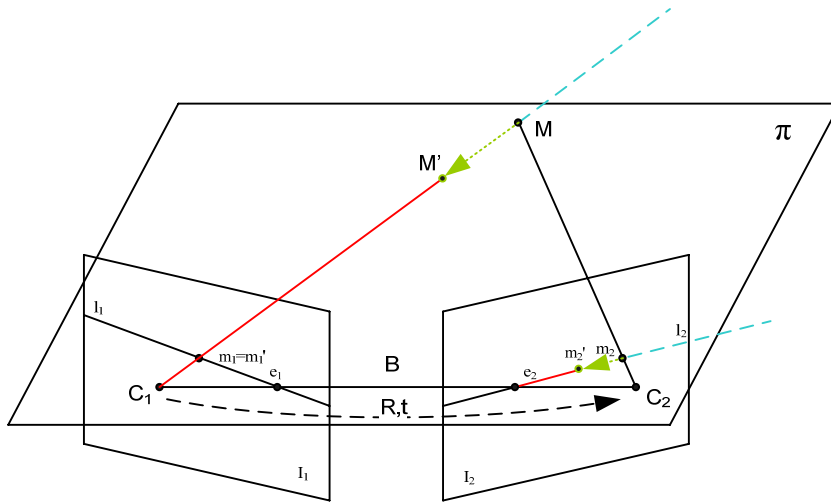


Abbildung 9: Zusammenhang zwischen Sichtstrahl und Epipolarlinie

Eine Verschiebung des Punkts  $M$  innerhalb der Epipolarebene, hat zur Folge, dass dieser entsprechend seiner Bewegung auf der zugehörigen Epipolarlinie wandert. Eine Epipolarlinie bildet somit den zu einem Punkt zugeordneten Sichtstrahl der anderen Kamera ab und beschreibt damit alle möglichen Korrespondenzpunkte.

Die Zusammenhänge werden durch die Epipolarbedingung formuliert. Diese besagt, dass ein korrespondierender Punkt der ersten Kameraabbildung nur auf der zugehörigen Epipolarlinie der zweiten Kameraabbildung liegen kann und umgekehrt. Die Suche nach einem korrespondierenden Punkt kann hierdurch auf eine Dimension reduziert werden. Der Zusammenhang ist mathematisch durch die Epipolargleichung gegeben.

$$\tilde{p}_2 = F \tilde{p}_1 = \tilde{p}_2 \bullet \tilde{l}_2 = 0 \quad (2-21)$$

Die Beziehung beider Kameras zueinander wird mit Hilfe der Transformationsmatrix  $F$  beschrieben. Diese definiert sich aus den intrinsischen Parametern beider Kameras und deren externen euklidischen Transformation.  $F$  wird Fundamentalmatrix genannt und beschreibt vollständig die Epipolargeometrie in Pixelkoordinaten. Die Fundamentalmatrix berücksichtigt die relative Orientierung, sowie die Eigenschaften beider Kameras.

$$F = K_2^{-T} [t]_{\times} R K_1^{-1} \quad (2-22)$$

Die Zusammenhänge und die Bestimmung der Fundamentalmatrix werden in der Literatur [ZHA98], [Schr05] ausreichend behandelt und können dort nachgelesen werden.

### **2.3.3 Korrespondenzanalyse**

Die Korrespondenzanalyse hat zum Ziel eindeutige, robuste und möglichst fehlerfreie Zuordnungen zwischen zwei Bildern zu finden. Zum Einsatz kommen pixelbasierte und merkmalsbasierte Verfahren.

Mithilfe von pixelbasierten Verfahren werden Bildpunkte zwischen zwei Bildern untersucht. Ein Bildpunkt zeichnet sich dabei durch seinen Farbwert aus. Da ein einzelner Bildpunkt für einen Vergleich nicht aussagekräftig genug ist, werden die umgebenen Bildstrukturen zur Analyse mit herangezogen. Korrespondenzen werden somit aufgrund von Pixelblöcken ermittelt. Solche Algorithmen werden daher auch Block-Matching Verfahren genannt.

Merkmalsbasierte Verfahren arbeiten mit den gleichen Algorithmen wie pixelbasierte Verfahren. Hierbei erfolgt aber nicht der Vergleich beliebiger oder aller Bildpunkte. In einem Vorverarbeitungsschritt werden über entsprechende Algorithmen interessante Kandidaten für die Korrespondenzanalyse bestimmt. Im Allgemeinen kommen Algorithmen zur Extraktion von Ecken, Kanten, Linien oder anderen Merkmalen zum Einsatz. Die Nutzung solcher markanten Bildpunkte ermöglicht bei der Korrespondenzanalyse eine eindeutigere Zuordnung zwischen den Bildmerkmalen und reduziert in der Regel den Verarbeitungsaufwand, da nicht alle Bildpunkte zur Analyse herangezogen werden.

Die Korrespondenzanalyse bringt durch die eingesetzten Verfahren auch einige Probleme mit sich. So ist es möglich, dass durch die geometrische Anordnung und fehlende Überlappung der Kameras eine Korrespondenz überhaupt nicht vorhanden ist. Zusätzlich fallen Korrespondenzen durch Verdeckungen innerhalb der Bilder heraus. Weitere Probleme ergeben sich durch periodische oder merkmalschwache Bildstrukturen, welche nicht mehr eindeutig zugeordnet werden können und aus der Analyse heraus fallen.



### 3 Analyse

#### 3.1 Bestehende Lösung

Diese Masterarbeit baut auf eine beim DLR angefertigten Diplomarbeit mit dem Thema „Feature Detection und Matching Verfahren zur Positions- und Lagebestimmung“ [ROS06] auf. Inhalt der Arbeit war es, aufgrund synthetischer Stereobildfolgen, verschiedene Methoden zur Extraktion von Punktmerkmalen zu untersuchen und mit den erfassten Punktmengen stabile Korrespondenzanalysen über mehrere Bildfolgen zu erzielen. Einzelne Korrespondenzanalysen erfolgen dabei zum einen zwischen den Bildern zweier Kameras eines gemeinsamen Zeitpunktes (Intra-Matching) und zum anderen zwischen den Bildern einer einzelnen Kamera zu den Zeitpunkten  $t_n$  und  $t_{n+1}$  (Inter-Matching). Das Ergebnis eines einzelnen Durchlaufs fasst Korrespondenzen aus zwei aufeinander folgenden Stereobildpaaren zusammen. Die gefundenen Punktmengen werden in Form von Ergebnislisten gespeichert. Die Gesamtanalyse ist durch folgenden Ablauf beschrieben:

1. Merkmalsextraktion und Erstellung einer Punktmenge aus Bild 1 des Zeitpunkts  $t_n$
2. Korrespondenzanalyse zwischen zwei Kamerabildern zum Zeitpunkt  $t_n$  (1. Intra-Matching)
3. Korrespondenzanalyse zwischen zwei Kamerabildern einer Kamera zu unterschiedlichen Zeitpunkten  $t_n$  und  $t_{n+1}$  (Inter Matching).
4. Korrespondenzanalyse zwischen zwei Kamerabildern zum Zeitpunkt  $t_{n+1}$  (2. Intra-Matching)

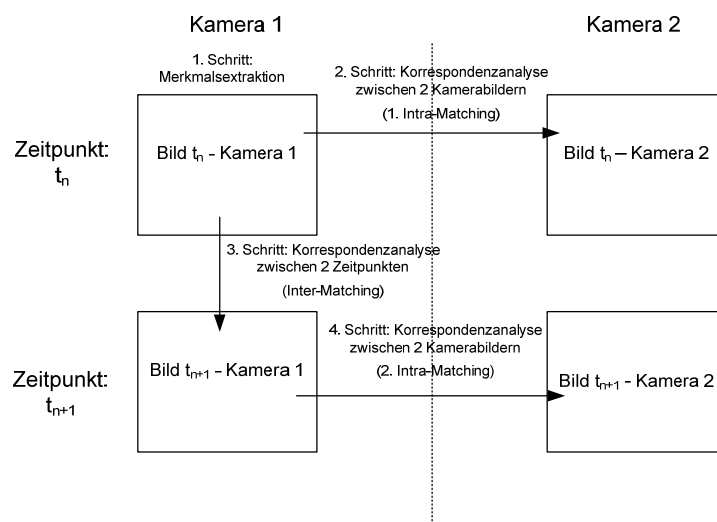


Abbildung 10: Korrespondenzanalyse über 2 Bildpaare

## **3.2 Programmfunktionalitäten**

Die vorliegende Testsoftware ist prototypisch erstellt worden und diente ursprünglich allein dem Ziel einen praktischen Nachweis der Ergebnisse der vorliegenden Diplomarbeit zu erbringen. Somit liegt keine Programmbeschreibung vor, welche detaillierte Rückschlüsse auf die Programmimplementierung zulässt. Die Ermittlung der Programmfunktionalitäten erfolgt daher anhand des vorliegenden Codes mit Hilfe der theoretischen beschriebenen Inhalte der Diplomarbeit.

### **3.2.1 Initialisierung**

Die vorhandene Lösung ermöglicht die Korrespondenzanalyse von Bildfolgen unter Verwendung verschiedener Algorithmen. Die Konfiguration dieser Algorithmen erfolgt über ein grafisches Eingabefenster oder über eine Konfigurationsdatei. Die übergebenen Konfigurationsdaten bestehen aus entsprechenden Parametern zur Initialisierung der Algorithmen zur Merkmalsextraktion und Korrespondenzanalyse, der Bildpfade und der Suchfenstergrößen.

### **3.2.2 Laden und Glättung der Bilddaten**

Für einen einzelnen Analysevorgang werden 4 Bilder geladen und anschließend geglättet. Dieser Vorgang dient zur Reduzierung des Rauschens um bessere Ergebnisse bei der Merkmalsextraktion zu erzielen. Als Filter kommen ein Binomialfilter oder ein Medianfilter zum Einsatz.

### **3.2.3 Merkmalsextraktion**

Durch Merkmalsextraktion erfolgt die Erstellung einer signifikanten Punktmenge für die Korrespondenzanalyse. Die Extraktion wird auf das erste der vier Bilder ausgeführt, wobei einer der implementierten Extraktionsalgorithmen zum Einsatz kommt. Alle Merkmalsextraktoren benötigen für den Einsatz den gleichen Umfang an Parametern, wobei diese lediglich in den Wertebereichen unterschieden sind. Die Bestimmung des Suchbereichs für ein Merkmal, erfolgt durch Aufteilung des Bildes in Kacheln. Durch Änderung der Kachelgrößen kann die Anzahl von extrahierten Merkmalen verändert werden. Der angewendete Algorithmus ermittelt für jede Kachel das beste Feature. Die Auswahl eines Features erfolgt basierend auf einem Schwellwert. Dieser Extraktionsschwellwert bestimmt dabei Qualitätsvorgabe des Merkmals und bildet das Grenzwertkriterium bei der Berechnung. Aus der vorangegangenen Arbeit wurden drei Operatoren untersucht und aufgrund guter Ergebnisse als nutzbar definiert:

- Kanade-Lucas-Tomasi (KLT),
- Moravec
- Harris Corner

Alle Ergebnisse der Merkmalsextraktion werden in Form einer Punktmenge aus x und y Koordinaten festgehalten. Diese dient als Basis zur weiteren Korrespondenzanalyse.

### 3.2.4 Korrespondenzanalyse durch Intra/Inter-Matching

Aufgrund der ermittelten Punktmenge aus dem ersten Bild, werden die korrespondierenden Punkte in den anderen Bildern gesucht. Zum Einsatz kommen Algorithmen sogenannter Blockmatching-Verfahren. Die Bestimmung von Korrespondenzen erfolgt über Pixelblöcke, die in zwei Bildern gebildet und aufgrund des eingesetzten Verfahrens auf Ähnlichkeitsmerkmale untersucht werden. Folgende Matchingalgorithmen sind bisher implementiert:

- Normierte Kreuzkorrelation (NCC)
- Mittlerer absoluter Fehler (SAD)
- Mittlerer quadratischer Fehler (SSD)
- Normierter Mittlerer quadratischer Fehler (NSSD)
- Census Transformation

Alle Punkte der Merkmalsliste werden nacheinander zur Korrespondenzanalyse herangezogen. Hierzu wird ein quadratisches Fenster um die Koordinaten des zu bearbeitenden Punkts im Ursprungsbild gebildet. Über seine Farbwerte wird versucht das Pattern eindeutig zu beschreiben.

32,65	32,66	32,67	32,68	32,69
33,65	33,66	33,67	33,68	33,69
34,65	34,66	34,67	34,68	34,69
35,65	35,66	35,67	35,68	35,69
36,65	36,66	36,67	36,68	36,69

Abbildung 11: Beispiel für ein 5x5 Pattern um das Pixel x:34 und y:67

Ist das Pattern des Ursprungsbilds beschrieben, wird der Suchbereich in Form der vordefinierten Größe um den analogen Punkt im Suchbild erzeugt. Sequentiell wird über jedes Pixel des Suchfensters ein Vergleichspattern gebildet und mithilfe des angewendeten Matching-Algorithmus versucht Ähnlichkeiten zwischen beiden Patterns zu ermitteln. Ein Pixel des Suchbildes gilt dabei als korrespondierend, wenn entsprechend des Algorithmus und seines eingestellten Schwellwerts, die größte Ähnlichkeit zwischen dem Quell- und Zielpattern besteht. Die ermittelten Koordinaten des korrespondierenden Pixels aus dem Suchbild werden den zugehörigen Pixelkoordinaten aus dem Ursprungsbild zugeordnet und zusammen abgespeichert.

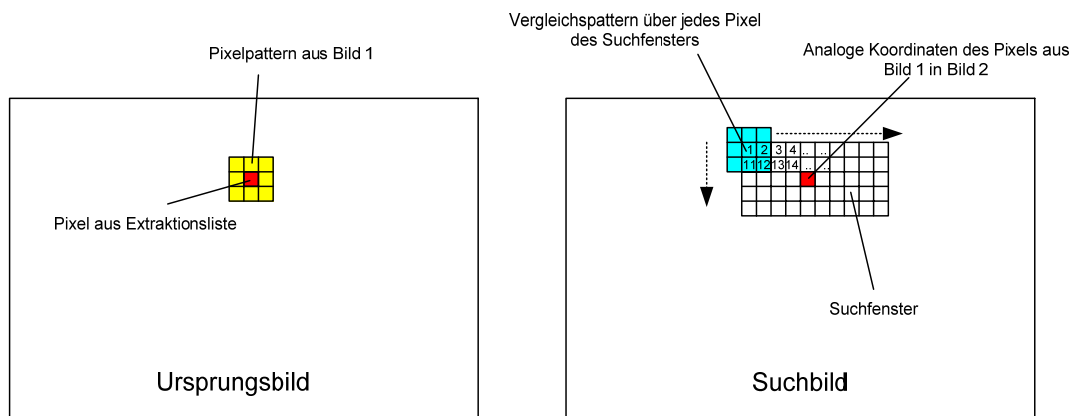


Abbildung 12: Matchingalgorithmus der vorliegenden Testsoftware

Der bisherige Algorithmus unterscheidet zwei verschiedene Arten der Korrespondenzanalyse, das Intra- und das Inter-Matching.

Das Intra-Matching beschreibt die Suche zwischen zwei Bildern desselben Zeitpunkts. Ausgangspunkt bilden zwei fest zu einander montierte Kameras. Lage- und Position beider Kameras sind zueinander konstant, wobei die Kameras unterschiedliche Eigenschaften haben können.

Das Inter-Matching betrachtet die Korrespondenzanalyse zwischen zwei Bildern einer Kamera zu unterschiedlichen Zeitpunkten. Die Kameraeigenschaften sind hierbei für beide Bilder konstant, da nur eine Kamera genutzt wird. Lage- und Position ist variabel.

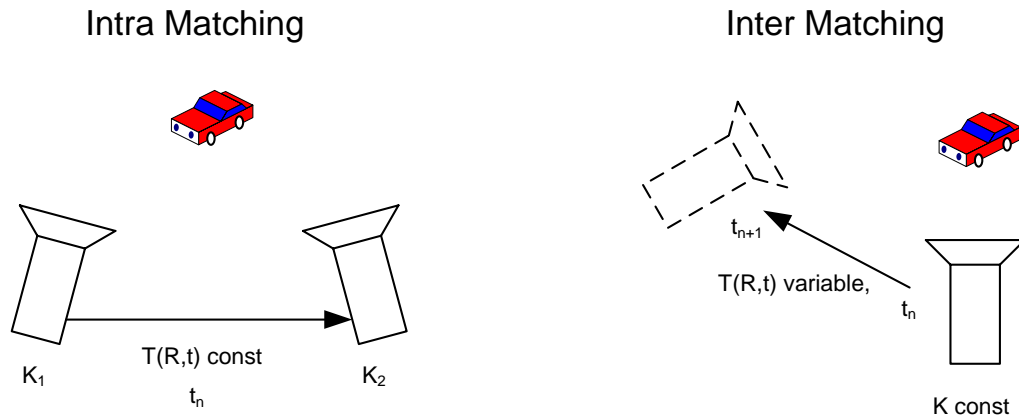


Abbildung 13: Klassifizierung des Intra-/Inter-Matching

### 3.2.5 Die Auflösungspyramide

Unter der Annahme, dass Merkmale in einem herunterskalierten Bild auch Merkmale in einem originalskalierten Bild sind, erfolgt der Einsatz einer sogenannten Auflösungspyramide. Hierbei wird die Suche von Merkmalen und deren Korrespondenzanalyse bei einem, um mehrere Skalierungsstufen, verkleinerten Bild durchgeführt. Somit verkleinern sich nicht nur der Bildbereich, sondern auch die Suchräume quadratisch, da weniger Pixel durchsucht werden müssen, um das gleiche Umfeld zu berücksichtigen. Nach einem kompletten Durchlauf der ersten Stufe werden die gefundenen Merkmale nur noch innerhalb eines sogenannten Refining-Fensters gematcht. Der Suchraum für ein Merkmal, welches in der nächst kleineren Stufe von den Koordinaten bekannt war, kann so in seinem direkten Umgebungsfeld in der nächst höheren Stufe wieder gematcht werden. Durch die Verkleinerung der Bilder und damit auch der Suchbereiche können Korrespondenzanalysen durch Reduzierung der zu durchsuchenden Pixel, beschleunigt werden. Trotz des Mehraufwands durch die Skalierungsstufen der Bilder reduziert sich gerade bei sehr großen Suchräumen der Suchaufwand enorm.

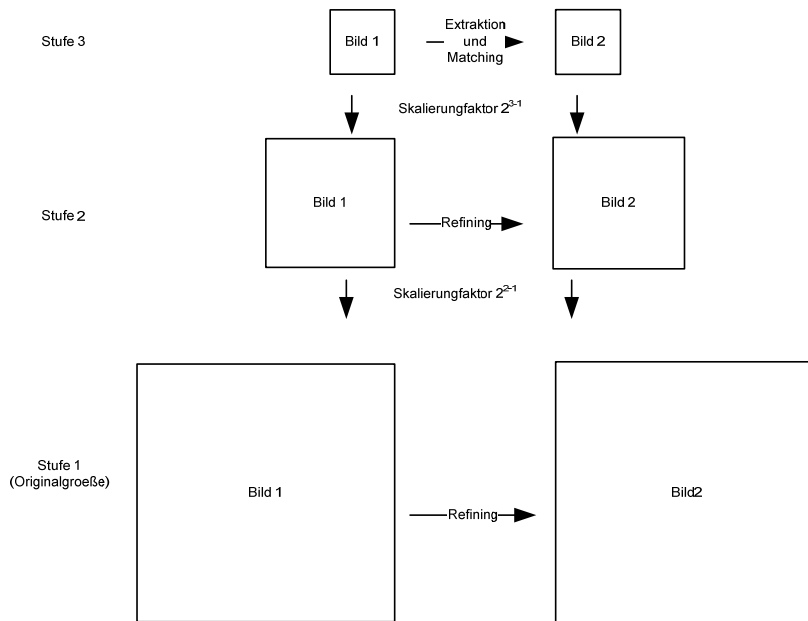


Abbildung 14: Beispiel für eine 3-stufige Auflösungspyramide zwischen 2 Bildern

### 3.2.6 Sonstige Funktionen

Eine weitere interessante Teilfunktion bildet das „Tracking over Pairs“. Der Ansatz verfolgt Merkmale über mehrere Bildpaare. Die Extraktion und die Korrespondenzanalyse eines ersten Bildpaars werden über mehrere Bildpaare beibehalten, wobei die Merkmalsmenge durch das Herausfallen von Punkten stetig abnimmt. Der Ansatz dient zur Beschleunigung der Durchlaufzeit, zeigt aber nur bei kleineren Bewegungen Wirkung, da hier Merkmale eine längere Lebensdauer haben.

### 3.2.7 Allgemeiner Programmdurchlauf

Aufgrund der analysierten Teilfunktionen kann jetzt eine allgemeine Softwarebeschreibung erstellt werden. Nach der Initialisierung der Parameter folgt das Laden der Bildpaare. Die Bildpaare werden auf die unterste Skalierungsebene des angegebenen Pyramidenlevels skaliert. Ist der Pyramidenlevel eins, erfolgt keine Skalierung.

Die Korrespondenzanalyse über vier Bilder erfolgt in vier Teilschritten. Begonnen wird mit der Merkmalsextraktion. Die Koordinaten der gefundenen Punkte werden als erste Ergebnisspalte in eine von drei Listen eingetragen.

Über ein erstes Intra-Matching werden im zweiten Schritt die Korrespondenzen zu den ermittelten Punkten im zweiten Kamerabild gesucht. Die gefundenen Punktkoordinaten bilden die zweite Spalte der Liste und werden den Ergebnissen aus dem ersten Bild direkt zugeordnet. Noch gültige Ursprungs- und Korrespondenzpunkte

sind als erste Ergebnisteilmenge in Form ihrer Koordinaten und eines Vergleichskoeffizienten beschrieben. Nicht wiedergefundene Punkte werden als ungültig gekennzeichnet und werden aus der Liste entfernt

Zur Korrespondenzanalyse zwischen dem ersten Bild des ersten Bildpaares (Zeitpunkt  $t$ ) und dem ersten Bild des zweiten Bildpaares (Zeitpunkt  $t+1$ ), werden die gültigen Koordinaten der ersten Ergebnisliste als neue Ausgangsmenge in eine leere zweite Liste kopiert. Das Inter-Matching zwischen den ersten Bildern beider Bildpaare bestimmt den dritten Schritt. Die hierbei wiedergefundenen Korrespondenzen werden den Koordinaten der zweiten Liste zugeordnet und gespeichert.

Der vierte und somit letzte Schritt umfasst das zweite Intra-Matching zwischen den Bildern des zweiten Bildpaars. Die gültigen korrespondierenden Bildkoordinaten des Inter Matching werden als neue Grundmenge in eine dritte Liste kopiert. Die Ergebnisse der Korrespondenzanalyse zwischen dem ersten und zweiten Bild des zweiten Paares bilden die zugehörigen Teilergebnisse der dritten Liste. Auch hier werden nicht wiedergefundenen Bildkoordinaten als ungültig markiert.

Um jetzt nur Punkte zu berücksichtigen, die in allen vier Bildern vorhanden sind, müssen die Ergebnisse der ersten beiden Listen, denen der dritten Liste angepasst werden. Durch einen Konsistenzcheck werden alle Ergebnisse der ersten und zweiten Liste auf ungültig gesetzt, wenn diese in der dritten Liste ebenfalls ungültig sind.

Bei Einsatz der Pyramide folgt das Refining zur Bestimmung der noch gültigen Koordinaten auf den nächst höheren Auflösungsstufen und es folgt die Speicherung der Ergebnisse. Hierzu werden die erste und letzte Ergebnisliste in Form eines Protokolls gespeichert. Beide Listen enthalten alle relevanten Koordinaten der Korrespondenzanalyse. Aus den Ergebnisprotokollen aller Durchläufe kann die spätere Berechnung der Bewegung erfolgen.

Ein Ergebnisprotokoll bezieht sich immer auf vier Bilder. Bei weiteren Durchläufen erfolgt die Weiternutzung des zweiten Bildpaars als erstes Bildpaar. Das nachfolgende Bildpaar der Testreihe wird als zweites Bildpaar geladen und die Korrespondenzanalyse erfolgt von neuem. Durch den Einsatz der Funktion „Tracking over Pairs“ wird die Merkmalsliste über mehrere Bildfolgen vorgehalten. Die Ergebnisse des zweiten Intra-Matching des letzten Durchlaufs werden als Grundlage des neuen Durchlaufs weiterverwendet. Die Merkmalsextraktion und das erste Intra-Matching entfallen somit bei diesen Durchläufen.

Der folgende Programmablaufplan beschreibt den Ablauf der vorhandenen Software. Er dient als Grundlage zur Erweiterung des Systems, wobei der Ablaufalgorithmus in seinen Grundzügen beibehalten werden kann.

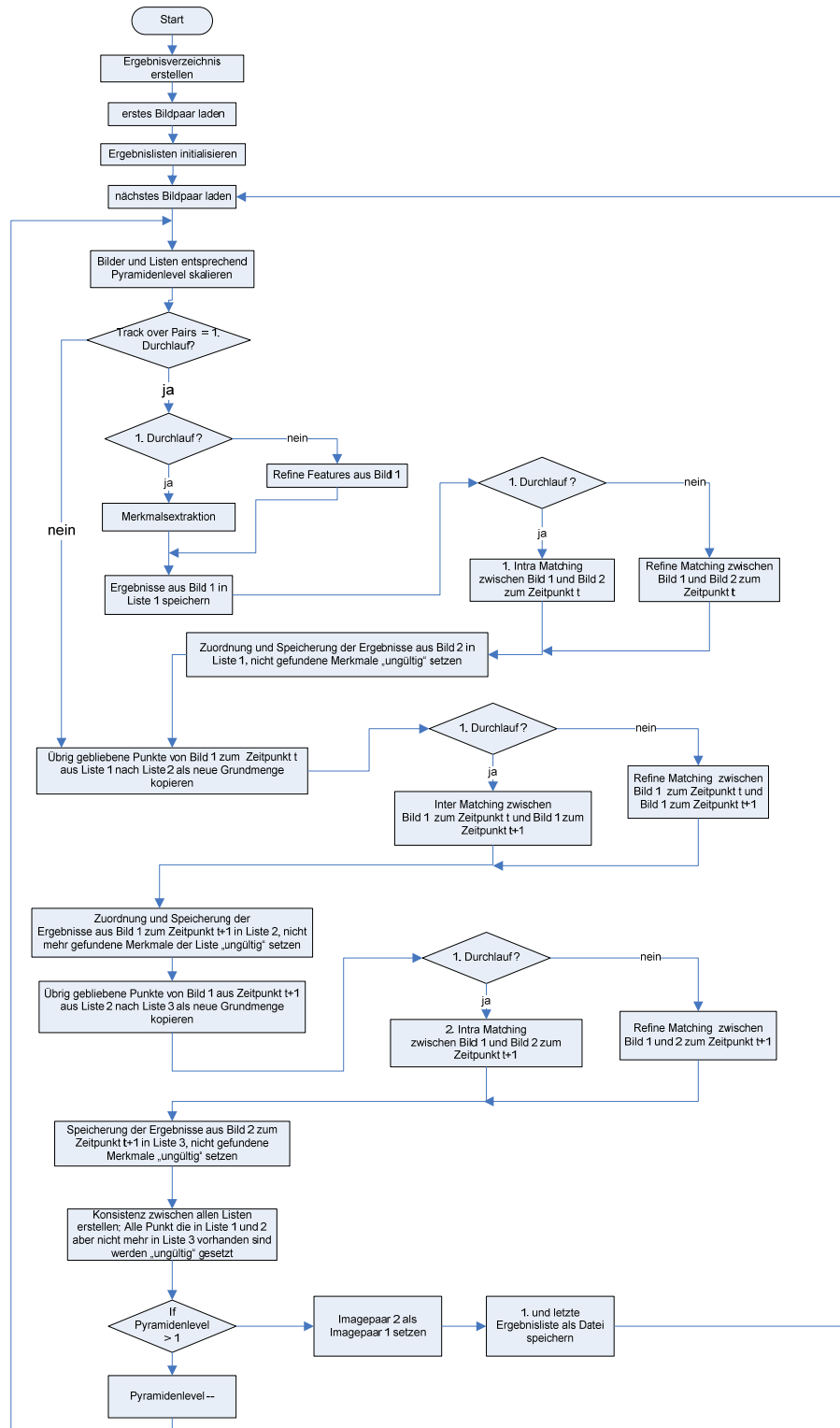


Abbildung 15: Programmablauf der Originalsoftware



### 3.3 Problemanalyse

Die Bestimmung von Suchräumen für die Korrespondenzanalyse erfolgt bisher, ohne Vorkenntnis über Kamerageometrie- und Lage. Die Größen der Suchbereiche werden zu Programmbeginn festgesetzt. Zum Zeitpunkt der Diplomarbeit lagen zum Nachweis der Ergebnisse synthetische Bildreihen aus entzerrten Bildern einer gleichförmigen Bewegung vor. Die Ergebnisse stützen sich dabei auf Suchbereiche von 891 bis 2121 Pixel pro Intra- oder Inter-Matching für Testfälle ohne Pyramide. Bei einer Größe für Vergleichspattern von 225 Pixeln, müssen zur Findung von Korrespondenzen zwischen 200475 und 477225 Pixel durchsucht werden. Durch den Einsatz der Pyramide kann die Rechenzeit bereits um ca. 70% vermindert werden. Die Klassifizierung der Ergebnisse erfolgte dabei in erster Linie anhand von Datensätzen mit konstanter Merkmalsanzahl und geringeren Durchlaufgeschwindigkeiten.

Als robuste Feature-Extraktoren werden KLT und Harris-Corner Operator beschrieben. Aufgrund des geringeren Rechenaufwandes brachte der Moravec allerdings bessere Durchlaufzeiten, wobei schlechtere Ergebnisse bei der Korrespondenzbestimmung das Resultat waren. Bei den Matching-Algorithmen erwiesen sich SAD und SSD als schnelle Algorithmen mit relativ guter Qualität. Die Metriken NCC und NSSD lieferten zwar ebenfalls gute Ergebnisse, dies jedoch bei schlechteren Laufzeitqualitäten.

Die größte Einschränkung innerhalb dieses Ansatzes ergibt sich aus den Angaben über die Größe des Suchfensters. Die Größe des Suchfensters wird über Parameterangaben frei gewählt. Hierbei ist die Größe von vornherein so anzupassen, dass alle möglichen Korrespondenzen der Bildfolgen durch das Suchfenster abgedeckt werden. Dieses Vorgehen setzt voraus, dass die maximale Bewegung der zu untersuchenden Bildfolge für das Intermatching und der Parallaxebereich für das Intramatching bekannt sind oder aufgrund von Erfahrungen geschätzt werden können. Um eine hohe Trefferwahrscheinlichkeit für Korrespondenzen zu erzielen, müssen relativ große Suchräume angesetzt werden. Zusätzlich steigt die Wahrscheinlichkeit für Fehlinterpretationen durch Ähnlichkeiten in der Bildstruktur. Diese Umstände verlangen einen erhöhten Rechen- und Zeitaufwand, den es zu reduzieren gilt.

## 4 Systementwurf

### 4.1 Allgemeiner Lösungsansatz

Aufgrund der Grundlagen der Stereoanalyse aus Kapitel 2.3 können durch Kenntnis über relative Lage und Eigenschaften zweier Kameraansichten Ähnlichkeitsbedingungen mit Hilfe der Epipolargeometrie formuliert werden. Intra- und Inter-Matching beruhen auf dem gleichen Matchingalgorithmus, wodurch die Epipolarbedingung unabhängig angesetzt werden kann. Der Unterschied ergibt sich jedoch aus der physikalischen Systembetrachtung, da dadurch verschiedene Datenquellen für die Berechnungen herangezogen werden müssen. Beim Intra-Matching ist die Rotation und Translation zwischen beiden Kameras konstant und braucht daher nur einmal ermittelt werden. Es ergeben sich jedoch in der Regel unterschiedliche Eigenschaften der Kameras, die zu berücksichtigen sind. Aufgrund der Abhängigkeit zur Zeitkomponente müssen beim Inter-Matching die Bewegungsdaten als variabel eingestuft werden. Rotation und Translation zwischen beiden Bildern ergeben sich aus der Bewegung der aufnehmenden Kamera.

Die Integration von Bewegungs- und Kameraeigenschaften in die Bildanalysen ermöglicht für einen Bildpunkt die Berechnung der Epipolarlinie in einem zweiten Bild. Die Ansätze der Epipolargeometrie würden helfen, die Suchräume für Korrespondenzen zu verkleinern und somit eindeutiger und schnellere Ergebnisse zu erzielen. Die bereits in der vorangegangenen Arbeit untersuchten und angewendeten Algorithmen zur Merkmalsextraktion und Korrespondenzanalyse können wiederverwendet werden. Die Konzepte der bestehenden Testsoftware für Korrespondenzanalysen soll für den realen Einsatz weiterentwickelt werden, so dass geometrische Abhängigkeiten des Kamerasystems und Nutzung von externen Bewegungsinformationen, berücksichtigt werden. Durch Auswertung und Nutzung der Daten wird versucht, Effizienz und Durchlaufzeiten des Systems bei gleichbleibender bzw. verbesserter Robustheit der Korrespondenzen erheblich zu verbessern. Bei Verwendung von realen Bildern, müssen zusätzlich die Verzeichnungseigenschaften berücksichtigt werden und in die Lösung mit einfließen. Die Weiternutzung von bereits gefundenen Merkmalen soll geprüft werden.

## 4.2 Systemaufbau

Zur Ermittlung von Lage- und Positionsdaten müssen die Systemkomponenten und ihre Bezugssysteme klar beschrieben und in Zusammenhang gebracht werden. Das zu untersuchende System fasst Daten aus zwei verschiedenen Sensoren zusammen. Hierbei ist der Aufbau eines Stereokamerasystems in Verbindung mit einem Messsystem zur Bereitstellung von Bewegungsdaten zu betrachten.

Bei Betrachtung des physikalischen Systemmodells ergibt sich die erste Teilkomponente in Form des Stereokamerasystems, welches aus zwei fest zueinander montierten Kameras besteht. Beide Kameras sind durch unterschiedliche Eigenschaften definiert, wobei Abstand und Lage zueinander konstant sind. Die Eigenschaften der Kameras lassen sich mit Hilfe von Kalibrationsprozessen bestimmen und über ihre intrinsischen Parameter in Form der Kameramatrix beschreiben. Die Orientierung und Lage einer Kamera auf ein Bezugssystem kann durch die extrinsischen Parameter angegeben werden. Messungen der Kameras werden in Form von Bilddaten bereitgestellt. Die zweite Teilkomponente ergibt sich durch Integration eines Systems zur Bewegungsmessung. Messdaten einer entsprechenden Einheit liefern Bewegungsdaten in Form einer Rotation und Translation. Da hier jedes beliebige System zur Bewegungsmessung herangezogen werden kann, soll auf eine Klassifizierung des Messsystems verzichtet werden. Die einfließenden Daten über Position und Lage beschreiben die Bewegung und Blickrichtungsänderung der aufnehmenden Kamera zwischen zwei Bildaufnahmen. Die Bewegung wird somit relativ zum letzten Aufnahmezeitpunkt beschrieben.

Zur Verkleinerung von Suchräumen beim Inter-Matching, sollen zugehörige Bewegungsdatensätze als Schätzung herangezogen werden. Die Bewegung aus Rotation und Translation ist hierbei über den Zeitpunkt der Messung definiert. Das Systemmodell basiert dabei auf der Annahme, das Bild und Bewegungsdaten beider Sensorsysteme synchron zum gleichen Zeitpunkt ermittelt werden können.

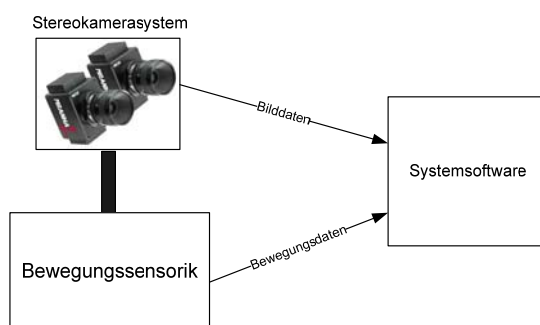


Abbildung 16: Systemmodell aus Stereokamerasystem und Bewegungssensorik

#### 4.2.1 Bezugssysteme ( $T_{tcp}$ , $T_{cam}$ , $T_{mov}$ )

Aufgrund der Zusammenhänge des euklidischen Raums sind Transformationen zwischen verschiedenen orthogonalen Koordinatensystemen definiert. Eine Transformation, bestehend aus Rotation und Translation kann somit in Form der linearen Transformationsmatrix  $T_E$  aus 2.1.5 beschrieben werden. Durch Bestimmung der Teilkomponenten des physikalischen Systemmodells können folgende Bezugssysteme identifiziert werden.

- Kamerasystem 1
- Kamerasystem 2
- Bewegungssensorik

Bedingt durch den Systemaufbau ist festzustellen, dass nur relative Messdaten auf das System Einfluss nehmen. Dadurch ergibt sich das Problem, dass die Bezugssysteme einzelner Systemkomponenten nicht über einen absoluten Punkt im Weltkoordinatensystem in Zusammenhang gebracht werden können. Dieser Umstand führt zu einer abgeschlossenen Betrachtung des Systems. Der Bezugspunkt des Gesamtsystems wird dazu in den Ursprung des Koordinatensystems von Kamera 1 gelegt.

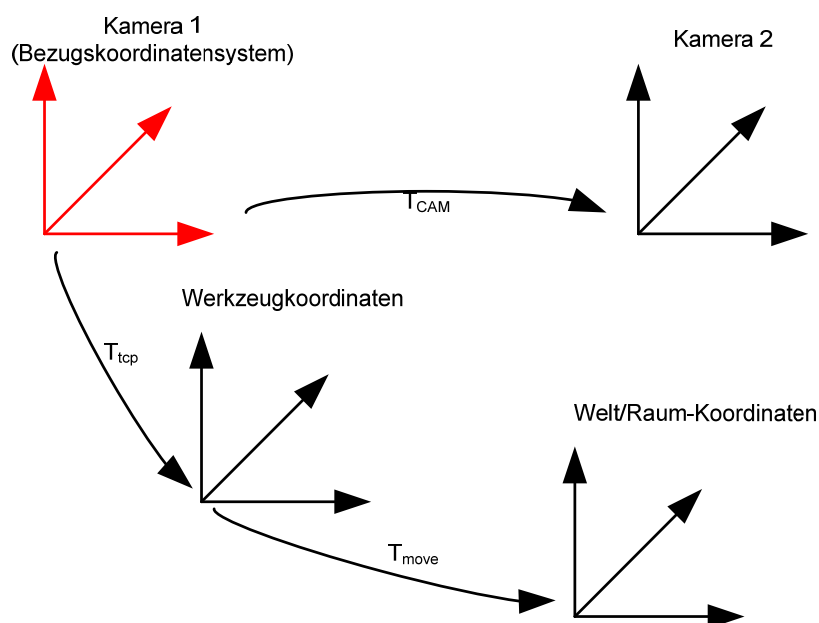


Abbildung 17: Zusammenhang zwischen den einzelnen Bezugssystemen

Die Einführung eines Werkzeugkoordinatensystems ermöglicht einen direkten Zusammenhang zwischen Bewegungssensor und Systemaufbau. Ein entsprechender Schritt ist allerdings nicht zwingend notwendig, da das Sensorsystem auch auf eines der Kamerasysteme bezogen werden könnte. Der Ursprung eines solchen Koordinatensystems wird durch den Tool-Center-Point (TCP) beschrieben. Dieser Begriff aus der Robotik beschreibt dabei einen Nichtsensorpunkt am physikalischen Systemaufbau, auf den Werkzeug- und Sensordaten z.B. zur Kalibration bezogen werden können. Der relative Abstand zwischen dem TCP und Systemaufbau bleibt dabei immer konstant, auch wenn das Gesamtsystem seine räumliche Lage ändert.

Die Bezugssysteme der Kameras sind durch die einzelnen Kamerakoordinatensysteme Kamera 1 und Kamera 2 definiert. Die Transformation  $T_{cam}$  beschreibt die Überführung von Bilddaten aus Kamera 1 in Kamera 2 und umgekehrt. Der Zusammenhang zwischen Kamera 1 und Werkzeugkoordinatensystem ist über  $T_{tcp}$  gegeben. Beide Transformationsmatrizen sind vom physikalischen Aufbau abhängig und können durch Kalibrationsmessungen ermittelt und definiert werden.  $T_{tcp}$  und  $T_{cam}$  sind innerhalb des Systemmodells konstant und beschreiben die relativen Bewegungsinformationen zwischen den Systemkomponenten. Messdaten des Bewegungssensors werden mit Hilfe des Werkzeugkoordinatensystems in Zusammenhang gebracht. Erfasste Messdatensätze über die Bewegung des Systems sind durch die Transformation  $T_{mov}$  zwischen Werkzeugkoordinatensystem und Welt/Raumkoordinaten beschrieben.

Durch die Gegebenheiten des physikalischen Systemmodells ergeben sich die geometrischen Zusammenhänge zwischen dem Intra- und Inter-Matching. Aufgrund der Verarbeitungspfade kann die resultierende extrinsische Orientierung  $dT$  zur Beschreibung der Position und Orientierung zwischen den Kameras durch die verwendeten Zusammenhänge bestimmt werden.

Das Intra-Matching beschreibt die Korrespondenzanalyse zwischen den Kameras eines fest montierten Stereosystems. Orientierung und Lage zwischen den Kameras kann über eine euklidische Transformation in Zusammenhang gesetzt werden. Als Bezugspunkt wird dabei der Ursprung des ersten Kamerasystems angenommen. Die relative Orientierung beider Kameras ist somit durch Transformationsmatrix  $T_{CAM}$  vollständig beschrieben, so dass für die extrinsische Orientierung  $dT$  folgendes gilt:

$$dT = T_{cam} \quad (4-1)$$

Das Inter-Matching beschreibt die Korrespondenzanalyse zwischen den Bildern einer Kamera zu unterschiedlichen Zeitpunkten. Da ohne weiteres kein geometrischer Zusammenhang zwischen der Orientierung der Kamera zu diesen Zeitpunkten herzustellen ist, sollen die Bewegungsinformationen der externen Sensorik unterstützend einfließen. Die relative Orientierung  $dT$  einer Kamera zu unterschiedlichen Zeitpunkten kann als Verkettung von euklidischen Transformationen  $T$  beschrieben werden. Somit ergibt sich für die Zeitpunkte  $t_n$  und  $t_{n+1}$  folgender Zusammenhang.

$$T_{(t_{n+1})} = T_{(t_n)} \bullet dT \quad (4-2)$$

Durch Umstellung ergibt sich für die relative Orientierung zwischen den Zeitpunkten

$$dT = T_{(t_{n+1})}^{-1} \bullet T_{(t_n)} \quad (4-3)$$

Da eine Korrespondenzanalyse sowohl zwischen den Bildern von Kamera 1 als auch zwischen den Bildern von Kamera 2 möglich ist, ergeben sich unterschiedliche Transformationspfade für die extrinsische Orientierung zwischen den Zeitpunkten. Unter Berücksichtigung von Lage und Richtung des definierten Koordinatenmodells können durch Einsetzen in  $T$  folgende Transformationsgleichungen erstellt werden:

Extrinsische Orientierung des Kamerasystem 1 zwischen den Zeitpunkten  $t$  und  $t+1$ :

$$dT = (T_{Move(t_{n+1})} \bullet T_{tcp})^{-1} \bullet T_{Move(t_n)} \bullet T_{tcp} \quad (4-4)$$

Extrinsische Orientierung des Kamerasystems 2 zwischen den Zeitpunkten  $t_n$  und  $t_{n+1}$ :

$$dT = (T_{Move(t_{n+1})} \bullet T_{tcp} \bullet T_{cam}^{-1})^{-1} \bullet T_{Move(t_n)} \bullet T_{tcp} \bullet T_{cam}^{-1} \quad (4-5)$$

Die Eigenschaften von Kameras sind über die intrinsischen Parameter definiert. Das Ergebnis liegt in Form der Kameramatrizen vor. Zur vollständigen Darstellung der Transformationsvorgänge werden im Systemmodell die Eigenschaften beider Kameras benötigt. Die Kameramatrizen  $K_1$  und  $K_2$  ermöglichen die Transformation von Kamerakoordinaten in Bildkoordinaten und umgekehrt. Die Transformation eines Objektpunkts in Bildkoordinaten ist in Formel 2-14 beschrieben.

### 4.3 Optimierung der Suchbereiche bei Matchingverfahren

Die Kenntnis über intrinsische Parameter der Kameras und deren relativen Orientierung und Lage, ermöglichen die Beschreibung des gesamten Transformationsprozess zwischen den Bildebenen zweier Kameras. Die grundlegende Bedingung zur Bestimmung von Korrespondenzen mit Hilfe der Epipolargeometrie ist somit erfüllt. Die Verkettung der Transformation ist in der Epipolargeometrie durch die Fundamentalmatrix  $F$  definiert. Aufgrund der verloren gegangenen Tiefeninformation innerhalb der Bildebene, ist eine punktgenaue Zuordnung allerdings nicht möglich. Das Ergebnis stellt sich in Form der Gradengleichung der Epipolarlinie dar, auf der alle möglichen Punkte des Sichtstrahls abgebildet sind.

#### 4.3.1 Berechnung der Epipolarlinie

Die Fundamentalmatrix kann auf zwei Wegen bestimmt werden. Zum einen kann die Berechnung mit Hilfe der Parameter über die äußere und innere Orientierung eines Stereokamerasystems erfolgen. Zum anderen kann die Fundamentalmatrix aus Punktmengen vorliegender Stereoansichten ohne Kenntnis der Kameraeigenschaften geschätzt werden. Für das vorliegende Systemmodell soll der erste Ansatz betrachtet werden. Aus den Grundlagen der Epipolargeometrie ergibt sich für die Epipolarlinie folgender Zusammenhang:

$$l_2 = F \bullet \tilde{p}_1 \quad (4-6)$$

Der Transformationsprozess fließt, wie in Formel 2-22 beschrieben, ein.  $K_1$  und  $K_2$  stellen dabei die Kameramatrizen dar und die Rotation wird durch die Matrix  $R$  beschrieben. Um die Multiplikation mit dem Translationsvektor  $t$  zu ermöglichen, wird dieser in Form einer schiefsymmetrischen Matrix  $[t_x]$  dargestellt. Somit ergibt sich die Matrixmultiplikation für die Fundamentalmatrix:

$$F_{3 \times 3} = [K_2^{-T}]_{3 \times 3} \bullet [t_x]_{3 \times 3} \bullet [R]_{3 \times 3} [K_1^{-1}]_{3 \times 3} \quad (4-7)$$

$$\text{mit } [t_x] = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} \quad \text{aus} \quad [t] = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Die Berechnung der Epipolarlinie für einen Bildpunkt kann jetzt, wie in Formel 4-2 beschrieben, umgesetzt werden, wobei der Ergebnisvektor als 3-dimensionaler Vektor  $l_2$  bis auf einen Skalierungsfaktor beschrieben ist.

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = [F]_{3 \times 3} \bullet \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (4-8)$$

Die Epipolarlinie ist durch die allgemeine Form  $[a,b,c]^T$  eines Richtungsvektor in der Ebene dargestellt und kann durch die allgemeine Gleichung einer Geraden in der Ebene beschrieben werden.

$$ax + by + c = 0 \quad (4-9)$$

Durch Umrechnung in  $f(x) = mx + d$  (4-10)

$$\text{mit} \quad m = -\frac{a}{b} \quad \text{und} \quad d = -\frac{c}{b} \quad \text{für} \quad b \neq 0$$

ergibt sich für die Epipolarlinie die Geradengleichung in Normalform.

**Eingrenzung des Bildbereichs:** Aufgrund des Wertebereichs der Geradengleichung für die Epipolarlinie ergeben sich für die Bestimmung der Suchbereiche einige Schwierigkeiten. Mit der Geradengleichung sind zwar alle möglichen Korrespondenzen zu einem Punkt beschrieben, allerdings liegt der Wertebereich der Gerade zwischen minus und plus unendlich. Somit ist es möglich, dass ein real korrespondierender Punkt außerhalb des Bildbereichs des zweiten Kamerasystems liegt und nicht durch dessen Bildebene erfasst ist.

Um hierbei den optimalen Ansatz zu entwickeln, müssen Maßnahmen zur Eingrenzung der Epipolarlinie auf den zweiten Bildbereich getroffen werden. Dazu muss ihre Geradengleichung auf die Größe der Bildebene eingeschränkt werden. Die Gleichung darf somit nur Pixel zwischen dem linken, rechten, oberen und unteren Bildrand berücksichtigen. Diese Anforderung führt zur Einführung von Grenzbereichen in u- und v-Richtung der Bildebene. Folgende Fälle müssen dabei berücksichtigt werden:



- Die Epipolarlinie liegt komplett außerhalb der Bildebene
- Die Geradengleichung wird in  $u$ -Richtung auf den linken und rechten Bildrand beschränkt
- Die Geradengleichung wird in  $v$ -Richtung auf den oberen und unteren Bildrand beschränkt

Die Berücksichtigung dieser Vorgaben erlaubt es, die Geradengleichung auf den gesamten Bildbereich zu begrenzen. Um eine Korrespondenz zu bestimmen, muss praktisch nur noch der Linienausschnitt durchsucht werden, der sich im Bereich der Bildebene befindet. Bei einem Bild von  $724 \times 724$  Pixel und einer Linienbreite von einem Pixel entspräche das dem Suchbereich einer Längsseite. Vergleicht man den Wert mit typischen Größen von angesetzten Suchfenstern des Ursprungssystems, ermöglicht der Ansatz bereits eine erhebliche Einschränkung des Suchbereichs.

**Korrektur der Epipolarlinie aufgrund von Pixelrasterung:** Das Einzeichnen bzw. Untersuchung einer Geraden innerhalb des Pixelrasters der Bildebene, erfordert elementare Korrekturmaßnahmen. Durch den entstandenen Rundungsfehler auf ganzzahlige Werte der Geraden kommt es je nach Steigung zu nicht berücksichtigten Zwischenräumen in eine der Koordinatenrichtungen. Die folgende Abbildung zeigt die Auswirkungen des Fehlers bei berechneten Werten in  $v$ -Richtung. Analog stellt sich das Problem bei der Berechnung von  $u$ -Werten.

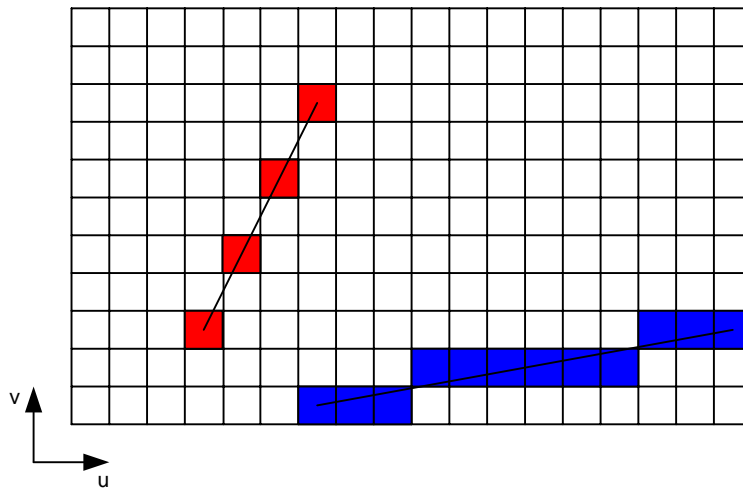


Abbildung 18: Abbildungsfehler bei Rasterung von Geraden (rot)

Die Berechnung von zugehörigen v-Koordinaten nach Formel 4-18 ergibt sich, durch die auf Pixel beschränkte Genauigkeit, nur für ganzzahlige u-Werte. Bedingt durch die Steigung wächst die Gerade in eine der Richtungen schneller an. Für berechnete v-Werte ergeben sich bei langsam ansteigender u-Richtung alle angenäherten Werte (blaue Linie). Wächst die Linie in v-Richtung schneller an, so ergeben sich für die berechneten Werte Lücken (rote Linie). Die Annäherung auf alle zu berücksichtigenden Pixel in v-Richtung ist durch den eingeschränkten Wertebereich in u-Richtung nicht gegeben.

Eine Lösung zur Behebung des Problems bietet der sogenannte Bresenham-Algorithmus. Dieser verfolgt die Idee, aufgrund der Steigung, iterativ die Setzrichtung für das nächste Pixel zu bestimmen. Grundsätzlich bestimmt sich der Wert eines Pixels in jedem Schritt durch Inkrementierung in langsamer Richtung. Durch Abprüfung einer Fehlervariablen wird dabei entschieden, ob ebenfalls eine Verschiebung in schneller Richtung erfolgt.

Die Fehlervariable selbst wird aufgrund des größeren Distanzwerts ( $d_u$ ) zwischen Start und Endpunkt gebildet. Bei jedem Schritt erfolgt die Subtraktion durch den Differenzwert des kürzeren Abstands ( $d_v$ ) und die Pixelkoordinate für die langsamere Richtung wird um eins erhöht. Unterschreitet  $d_u$  dabei den Nullwert, erfolgt die zusätzliche Inkrementierung der Koordinate in schnellerer Richtung. Der Ursprungswert von  $d_u$  wird danach auf den aktuellen Wert der Fehlervariable aufaddiert und somit für die nächste Stufe neu initialisiert.

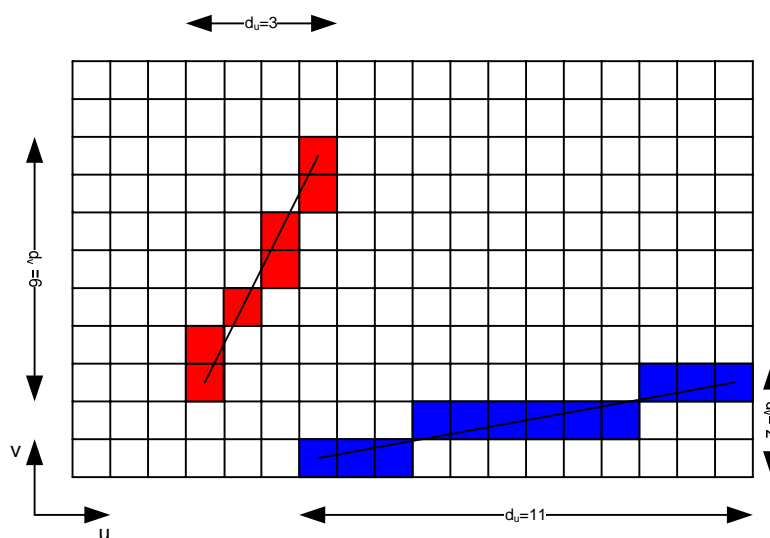


Abbildung 19: Annäherung von Pixelgeraden nach Bresenham

### 4.3.2 Einschränkung des Tiefenbereichs

Aufgrund der Kenntnis der intrinsischen und extrinsischen Eigenschaften des gesamten Stereokamerasystems lässt sich ein Bildpunkt von einer Bildebene in eine andere Bildebene transformieren. Durch die fehlende Tiefeninformation stellt sich der transformierte Punkt in der zweiten Bildebene nur noch als Epipolarlinie über den gesamten Bildbereich dar. Durch Definition des Tiefenbereichs soll der Wertebereich der Epipolarlinie definiert und diese dadurch weiter eingeschränkt werden.

Durch den geometrisch bedingten Aufbau der Kameras kann es sein, dass der Brennpunkt  $C$  der einen Kamera nicht in der Bildebene der anderen Kamera abgebildet wird. Somit liegt der Epipol für die Quellkamera außerhalb des Bildbereichs der Zielkamera. Korrespondenzen die auf der Epipolarlinie zwischen dem Epipol und dem Beginn der Bildebene liegen, existieren somit in der zweiten Bildebene gar nicht. Betroffen hiervon sind Objektpunkte, die nah zum ersten Kamerasystem stehen und durch das zweite Kamerasystem nicht erfasst werden können.

Ein ähnliches Problem ergibt sich für weit entfernte Objektpunkte, die zur Korrespondenzanalyse herangezogen werden. Auch diese können eventuell durch fehlende Überlappung nicht mehr von der Zielkamera erfasst werden. Dieser Zustand beschreibt alle möglichen Korrespondenzen auf der Epipolarlinie die sich außerhalb des Bildbereichs der Zielkamera gegen unendlich befinden.

Ein Objektpunkt im Kamerakoordinatensystem, der durch Transformation aus der Bildebene gewonnen wurde, ist für die dritte Ebene durch die fehlende Tiefeninformation nur noch durch den Skalierungsfaktor beschrieben. Somit fehlt die Information, in welchen Abstand der Punkt zur Kamera steht. Definiert man nun künstlich einen minimalen und maximalen Abstand, können die daraus resultierenden 3D-Punkte, direkt in die Bildebene des zweiten Kamerasystems transformiert werden. Beide Punkte liegen dabei im Wertebereich der Epipolarlinie des selektierten Bildpunkts und ermöglichen diese auf einen Start- und Endwert einzuschränken.

Die hinreichende Kenntnis über den zu untersuchenden Tiefenbereich einer Bildszena ermöglicht damit eine weitere Einschränkung des Suchbereichs. Durch Hinzufügen von definierten Tiefeninformationen zu einem Bildpunkt kann die Epipolarlinie auf einen kleineren Ausschnitt reduziert werden. Es werden dadurch nur noch die möglichen Korrespondenzen beschrieben, die sich innerhalb des vorgegebenen Tiefenbereichs befinden. Der Tiefenbereich sollte daher so gewählt sein, dass er sich mit großer Wahrscheinlichkeit in der Bildebene der zweiten Kamera abbilden lässt. Der

vollständige Algorithmus zur Bestimmung der eingeschränkten Epipolarlinie kann wie folgt beschrieben werden:

1. Berechnung der Matrix  $H$  zur Transformation eines Punktes aus den Objektkoordinaten des ersten Kamerasystems in die Bildkoordinaten des zweiten Kamerasystems

$$[H]_{3 \times 4} = [K_2 \mid 0]_{4 \times 3} \bullet [dT]_{4 \times 4} \quad (4-11)$$

1. Transformation des Punktes  $P1 (u,v)$  aus der Bildebene in das Kamerakoordinatensystem der ersten Kamera

$$\begin{bmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{bmatrix} = [K_1^{-1}]_{3 \times 3} \bullet \begin{bmatrix} u_{P1} \\ v_{P1} \\ 1 \end{bmatrix} \quad (4-12)$$

2. Objektpunkts mit minimaler Tiefeninformation  $d_{\min}$  für Objektkoordinaten des zweiten Kamerasystems ermitteln

$$\begin{bmatrix} X_{P \min} \\ Y_{P \min} \\ Z_{P \min} \end{bmatrix} = [H]_{3 \times 4} \bullet \begin{bmatrix} x_{c1} \bullet d_{\min} \\ y_{c1} \bullet d_{\min} \\ z_{c1} \bullet d_{\min} \\ 1 \end{bmatrix} \quad \text{mit } Z_{c1} = 1 \quad (4-13)$$

3. Objektpunkt mit maximaler Tiefeninformation  $d_{\max}$  für Objektkoordinaten des zweiten Kamerasystem ermitteln

$$\begin{bmatrix} X_{P \max} \\ Y_{P \max} \\ Z_{P \max} \end{bmatrix} = [H]_{3 \times 4} \bullet \begin{bmatrix} x_{c1} \bullet d_{\max} \\ y_{c1} \bullet d_{\max} \\ z_{c1} \bullet d_{\max} \\ 1 \end{bmatrix} \quad \text{mit } Z_{c1} = 1 \quad (4-14)$$

4. Berechnung der Start und Endpunkte der Epipolarlinie für die Bildebene im zweiten Kamerasystem

$$\text{Startwert: } P_{c2 \min} = \begin{bmatrix} u_{\min} \\ v_{\min} \end{bmatrix} = \begin{bmatrix} \frac{x_{P \min}}{z_{P \min}} \\ \frac{y_{P \min}}{z_{P \min}} \end{bmatrix} \quad (4-15)$$

$$\text{Endwert: } P_{c2 \max} = \begin{bmatrix} u_{\max} \\ v_{\max} \end{bmatrix} = \begin{bmatrix} \frac{x_{P \max}}{z_{P \max}} \\ \frac{y_{P \max}}{z_{P \max}} \end{bmatrix} \quad (4-16)$$

Somit ergibt sich ein begrenzter 1-dimensionaler Suchbereich aufgrund des berechneten Start- und Endwerts in der Bildebene des zweiten Kamerasystems. Trotz des eingeschränkten Suchbereichs ist es nicht ausgeschlossen, dass der Wertebereich außerhalb der Bildebene liegt. Somit muss auch dieser auf Gültigkeit geprüft und entsprechend angepasst werden. Zur Ermittlung der Punkte der Bildaußenbereiche, kann die Geradengleichung der Epipolarlinie in Normalform bestimmt und auf entsprechende Grenzwerte angewendet werden.

Geradengleichung:

$$v = mu + b \quad (4-17)$$

$$\text{mit } m = \left( \frac{u_{\max} - u_{\min}}{v_{\max} - v_{\min}} \right) \quad \text{und} \quad b = v_{\min} - m \bullet u_{\min} = v_{\max} - m \bullet u_{\max}$$

### 4.3.3 Nutzung der Ansätze der Epipolargeometrie

**Parametrierung:** Die Eingrenzung der Epipolarlinie erfordert weitere Parametrierungen bei der Initialisierung des Matching-Algorithmus. Neben dem Korrelationsschwellwert und der Größe des Korrelationsfensters ist die Angabe der minimalen und maximalen Distanz erforderlich. Ungenaue Bewegungsdaten oder Kameraeigenschaften, können eine Verbreiterung des Suchbereichs beim Intra- und Inter-Matching erforderlich machen. Im Gegensatz zum ersten Entwicklungsansatz müssen erweiterte Systemeigenschaften aus der Kamerakalibration und einer Bewegungssensorik integriert werden. Zur Bestimmung der Transformationsdaten werden für jede Kamera die Kameramatrix, die relativen Orientierungen zwischen den Ka-



#### 4.4 Weiternutzung gefundener Merkmale

Die Wiederverwendung bereits gefundener Merkmale verspricht weitere Möglichkeiten zu Einsparung von Verarbeitungszeit. Da das zweite Bildpaar eines vorangegangenen Analysevorgangs zum ersten Bildpaar des nächsten Durchlaufs wird, können bereits vorhandene Features und Korrespondenzen gehalten und somit wiederverwendet werden. Die direkte Wiederverwendung der Ergebnisliste aus dem letzten Durchlauf ist allerdings nicht möglich, da erst eine Neuordnung der Merkmale zu den entsprechenden Bildbereichen erfolgen muss.

Die Zuordnung von Features zu einzelnen Bildbereichen erfolgt aufgrund der Bildkachelung. Jeder Kachel darf ein Feature zugeordnet werden. Durch die Bewegung der Bilder ändern sich jedoch die Positionen der Merkmale. Bei Wiederverwendung müssen diese daher den Kacheln neu zugeordnet werden. Die Möglichkeit, dass mehrere Features der vorhandenen Ergebnisliste derselben Kachel zugeordnet werden, wird durch die alleinige Nutzung des besten Features umgangen. Schlechtere Features einer Kachel entfallen somit bei der Wiederverwendung.

##### 4.4.1 Nutzung der Merkmalswiederverwendung

Die Weiternutzung bereits gefundener Features wird in den vorhandenen Ansatz zur Merkmalsextraktion integriert. Der im ersten Systementwurf entwickelte Ansatz „Tracking over Pairs“ zur Weiternutzung von Features über mehrere Bildfolgen entfällt und wird durch die neue Lösung ersetzt.

Die Wiederverwendung von Merkmalen basiert auf den Ergebnissen des letzten Intra-Matchings eines vorangegangenen Durchlaufs. Die dort untersuchten Bilder bilden im neuen Durchlauf das erste Bildpaar. Somit werden die weiterhin gültigen Punkte als neue Ausgangsmenge angenommen. Es folgt die Anpassung an die neuen Gegebenheiten. Die Ergebnismenge muss bei Einsatz der Pyramide entsprechend der Bilder skaliert werden. Ist der Maßstab der Bilder und die Ergebnismenge zueinander konform, folgt die Zuordnung der noch gültigen Features zu den Bildkacheln. Aufgrund der Koordinaten des gültigen Features wird die Bildkachel berechnet. Der Ort einer Bildkachel ist aufgrund der Größe über einen Spalten- und Zeilenindex im Bildbereich beschrieben

$$Spalte_{Kachel} = \frac{u}{Kachell\ddot{a}nge} \quad Zeile_{Kachel} = \frac{v}{Kachelbreite} \quad (4-18)$$

Der Kachelindex selbst ergibt sich aus den abgeschnittenen ganzzahligen Anteilen der Spalten- und Zeilenwerte und kann durch Kenntnis der Kachelanzahl errechnet werden.

$$ID = (Zeile \in N) * \frac{Kachelanzahl}{Zeile} + (Spalte \in N) \quad (4-19)$$

Ein Feature wird aufgrund dieser ID der richtigen Kachel zugeordnet. Beinhaltet eine Kachel bereits ein Feature, erfolgt die Prüfung anhand des besseren Koeffizienten. Das stärkere Feature wird dabei der Kachel zugeordnet und das Schwächere verworfen. Aufgrund der neu geordneten Liste wird die Merkmalsextraktion nur auf Kacheln durchgeführt die noch kein gültiges Feature enthalten.

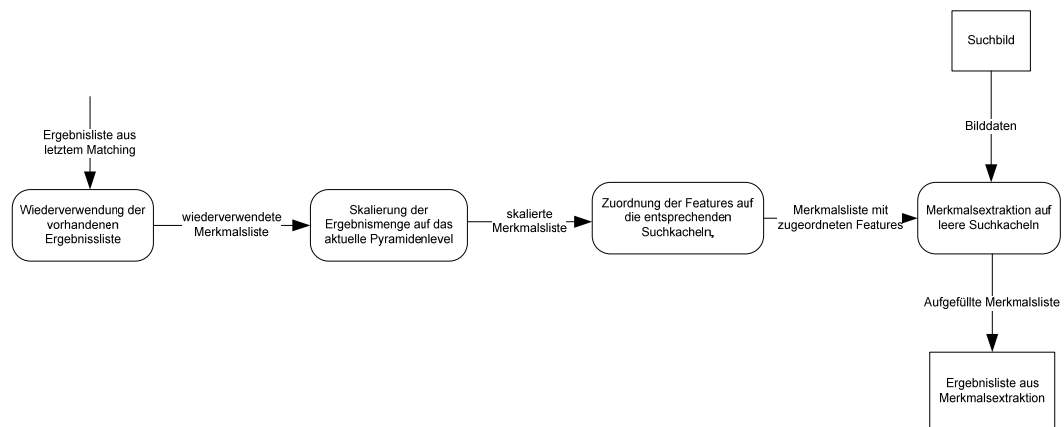


Abbildung 21: Wiederverwendung von Merkmalen

#### 4.5 Alternativer Matchingansatz durch 2. Inter-Matching

Für die Bewegungsanalyse werden generell 2 Bildpaare eines Stereokamerasystems herangezogen. Aufgrund der Einschränkung der Epipolargeometrie und der Nutzung externer Bewegungsdaten besteht die Möglichkeit, dass der Ansatz erstes Intra-Matching, Inter Matching und zweites Intra-Matching nicht mehr optimal ist. So kommt es bei kleineren Bewegungen. Gerade bei kleinen Bewegungen, bei denen die relative Lage kleiner ist als die der Kameras, können durch das Inter-Matching kleinere Suchräume erzeugt werden. Somit scheint es in diesen Fällen sinnvoller, mit Hilfe der externen Bewegung, ein zweites Inter-Matching über die Bilder der zweiten Kamera anzusetzen. Die erweiterte Nutzung der Epipolargeometrie auch beim Inter-Matching, ermöglicht flexiblere Ansätze bei der Gestaltung des Gesamtvorgangs. Durch Nutzung dieser Kenntnis kann der Arbeitsaufwand durch Optimierung des letzten Matching-Schritt weiter reduziert werden.



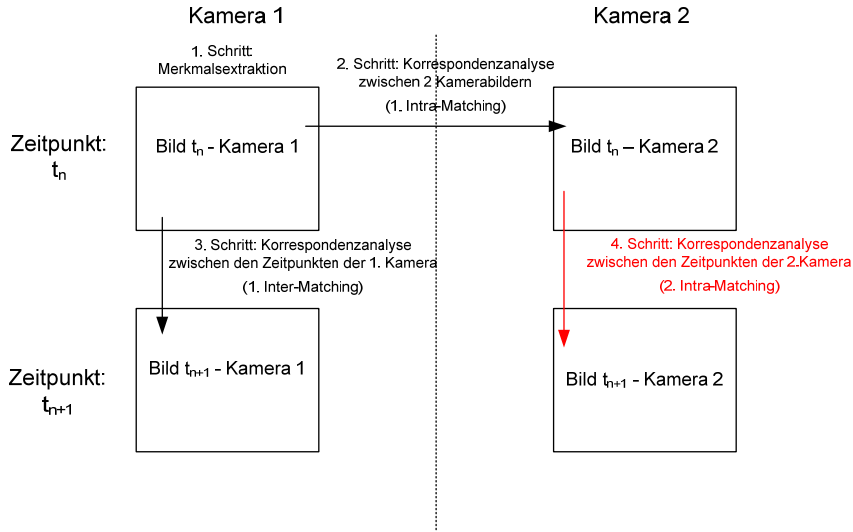


Abbildung 22: Matchingalternative durch zweites Inter-Matching

#### 4.6 Verzeichnungskorrektur bei realen Bilddaten

Reale Bilddaten eines Kamerasystems enthalten in der Regel nichtlineare Verzeichnungen. Die bisher verfolgten Ansätze zur Erstellung und Einschränkung der Suchbereiche mit Hilfe der Epipolargeometrie basieren jedoch auf linearer Bildgeometrien. Somit ist es nicht möglich, diese direkt auf die verzeichneten Bilder anzuwenden. Zur Behebung des Problems muss der Verzeichnungsfehler kompensiert werden.

Die Epipolarlinie beschreibt mögliche Korrespondenzen eines unverzeichneten Bildes in Form einer Geraden, da die Berechnung der Suchbereiche aufgrund der linearen Kamerageometrie basiert. Um den Algorithmus auf gefundene Merkmale eines nichtlinearen Bildes anzusetzen, müssen deren Koordinaten in ein lineares Bildsystem überführt werden.

$$\delta_u = u(K_1 r^2 + K_2 r^4 + K_3 r^6) \quad (4-20)$$

$$\delta_v = v(K_1 r^2 + K_2 r^4 + K_3 r^6) \quad (4-21)$$

$$\text{mit } r = \sqrt{u^2 + v^2} \quad (4-22)$$

Aufgrund dieser Transformation ist die Bestimmung des Suchbereichs mit Hilfe der Epipolarlinie auch für verzeichnete Pixel möglich. Allerdings liegen die Pixel des Suchbereichs ebenfalls als unverzeichnete Bilddaten vor. Zur Nutzung des Matching-

Algorithmen müssen diese daher mit den Originalbilddaten wieder in Zusammenhang gebracht werden.

Korrespondenzanalysen werden aufgrund von Vergleichsrechnungen auf Pixelpattern durchgeführt. Dieser Ausgangspunkt zieht eine nähere Betrachtung auf die verbundene Anzahl an Transformationen zwischen verzeichneten und unverzeichneten Bilddaten nach sich. Ein Matching-Vorgang im linearen Bereich würde die Anzahl an Transformationen durch die zusätzlichen Vergleichspattern erheblich erhöhen. Deshalb wird hierbei von einer Verarbeitung im linearen Bereich abgesehen. Es folgt daher die Transformation der Suchbereichspixel in den nichtlinearen Bereich.

Durch die Abhängigkeit zwischen Bildkoordinaten und Verzeichnisfehler kann die Verzeichnung eines Punkts nur durch Annäherung über ein iteratives Verfahren bestimmt werden. Mit Hilfe der Fixpunktiteration können die Formeln 4-18 bis 4-19 auch in Gegenrichtung angewendet werden. Der nicht verzeichnete Punkt dient dabei als Ausgangsvariable. Es folgt für ihn die Berechnung der Verzeichnung, welche aufaddiert wird. Durch Differenzbildung zwischen dem alten und dem neu errechneten Wert wird die Abweichung bestimmt. Diese wird mit einem Fehlerschwellwert  $\varepsilon$  verglichen. Ist die Fehlerabweichung größer, folgt die Berechnung der Verzeichnung für den neuen Punkt und der Schritt der Differenzbildung wird wiederholt. Das Verfahren wird beendet, wenn der reale Fehler den Schwellwert unterschreitet oder wegen nicht vorhandener Konvergenz nach  $n$  definierten Schritten abgebrochen wird.

Für das Abbruchkriterium gilt:

$$|P(u, v)_{n+1} - P(u, v)_n| \leq \varepsilon \bullet |P(u, v)_{n+1}| \quad (4-23)$$

#### 4.7 Einflüsse auf den Pyramidenansatz

Die Bestimmung kleinerer Suchbereiche wirft die Frage nach der Effektivität des Pyramidenansatzes auf. So ist es unumstritten, dass bei sehr groß angesetzten Suchbereichen, die Skalierung auf einen kleineren Bildbereich, den Einsatz der Pyramide am effizientesten macht. Durch die Bildverkleinerung werden ebenfalls die Suchbereiche verkleinert. Zur Bestimmung und Verarbeitung der Features werden in der Regel weitaus weniger Pixel herangezogen, was somit eine schnellere Extraktion und Korrespondenzanalyse ermöglicht. Die Summierung der verarbeiteten Pixel des Suchbereichs der höchsten Skalierungsstufe mit denen der Refining Bereiche aus den anderen Pyramidenstufen liegt dabei niedriger als der ursprünglich angesetzte Suchbereich ohne Pyramide. Die mit Hilfe dieses Ansatzes erreichte Ersparnis ist jedoch Abhängig von der Größe des angesetzten Suchbereichs.

Es wird angenommen, dass ein ermittelter Suchbereich 100 Pixel umfasst. Somit kann dieser durch den eindimensionalen Suchbereich, bei einer 3-stufigen Pyramide in der untersten Stufe auf 25 Pixel eingegrenzt werden. Bei einem typischen Refining-Fenster von 25 Pixeln ergibt das insgesamt 75 Pixel die insgesamt verarbeitet werden. Das ergibt eine Arbeits- und Zeitersparnis von 25%. Wie verhält sich der Ansatz jedoch bei noch kleineren Suchbereichen? Ist ein Suchbereich von nur 25 Pixeln ermittelt, ergeben sich bei gleichen Randbedingungen nach Durchlauf des Refining 56 verarbeitete Pixel. Damit erhöhen sich für dieses Beispiel der Arbeitsaufwand und die Verarbeitungszeit um ca. 56 %.

Aufgrund der relativen Orientierung zwischen zwei Bildern, kann es zu sehr kurzen eingeschränkten Epipolarlinien kommen, so dass der Pyramidenansatz sich hierbei negativ auf die Verarbeitungszeit auswirken kann. Es ist daher mit zu rechnen, dass der Pyramidenansatz aufgrund von Kameraaufstellung und kleinen Bewegungen, sowohl beim Intra-Matching als auch beim Inter-Matching, längere Verarbeitungszeiten verursacht. Für die Merkmalsextraktion bleiben die Auswirkungen der Pyramide in jedem Fall erhalten, da das Ergebnis von der Bildgröße abhängt.

## 5 Testsoftware

Ein Nachweis über die Auswirkungen der einzelnen Optimierungsansätze erfolgt durch Modifizierung der bereits vorhandenen Testsoftware. Diese wird soweit verändert, dass die in dieser Arbeit beschriebenen Ansätze vollständig umgesetzt sind. Hierzu sind Veränderungen in der Algorithmusstruktur, sowie in der Klassenhierarchie notwendig. Systemschnittstellen müssen untersucht und entsprechend erweitert werden. Die verwendeten Operatoren zur Extraktion und Korrespondenzanalyse können wiederverwendet werden. Da eine Vergleichsanalyse zwischen den einzelnen Operatoren nicht Inhalt dieser Arbeit ist, kann der Einsatz auf einzelne Lösungsansätze der Matchingverfahren reduziert werden. Zum Einsatz kommen die Merkmalsextraktoren KLT und Moravec und die Matching-Algorithmen NCC, SAD und SSD. Im Weiteren sollen grundlegende Veränderungen und Umsetzungen der Testsoftware vorgestellt und näher beschrieben werden.

### 5.1 Systemschnittstellen

#### **Funktionale Parameter**

Zur Nutzung der Analysesoftware sind allgemeine Initialisierungsparameter für die Funktionalität erforderlich. Die Bereitstellung der Daten erfolgt wie bisher über eine Textdatei. Als Funktionsparameter werden alle Daten zusammengefasst die zum Betrieb der Software nötig sind. Hierzu gehören Eingabedaten zur Auswahl der Merkmalsextraktions- und Matchingalgorithmen, so wie Umgebungsinformationen zu Quell- und Zieldaten. Alle Parameter werden innerhalb der Textdatei in Form einer Parameterbeschreibung mit zugehörigem Wert eingetragen. Die Datei dient als eigentliche Benutzerschnittstelle zur Bedienung der Software.

#### **Integration der Kameraeigenschaften**

Die Kameraeigenschaften des Stereosystems werden über eine Kalibrationsdatei bereitgestellt. Die Datei beinhaltet extrinsische und intrinsische Parameter, sowie die Verzeichnungseigenschaften der Kameras des eingesetzten Stereosystems. Die extrinsischen Parameter setzen sich aus den relativen Orientierungen der Kameras in Bezug auf das Werkzeugkoordinatensystem zusammen.

#### **Bilddaten aus dem Kamerasystem**

Bilddateien simulieren die Bereitstellung von Bilddaten aus dem konfigurierten Kamerasystem. In einer realen Anwendung entspräche das der Systemschnittstelle zu einem externen Kamerasystem. Die zu verarbeitenden Bilddaten fließen ausschließ-

lich in Dateiformat ein. Die untersuchenden Bildfolgen stehen dabei im direkten Zusammenhang zu den Kameraeigenschaften.

### Externes Sensorsystem

Die Schnittstelle beschreibt die Einspeisung von Bewegungsdaten aus dem externen Sensorsystem. Die Bewegungsdaten sind vorverarbeitet und liegen als Datei in Form von Translations- und Rotationsdaten vor. Jeder Bewegungsdatensatz innerhalb der Datei steht in direktem Bezug zum Aufnahmezeitpunkt eines Bildpaars der Kameraschnittstelle.

### Ausgabedaten

Zu jeden Matching-Durchlauf gibt ein Ausgabeprotokoll Auskunft über die gefundenen Korrespondenzen der vier untersuchten Bilder. Das Ausgabeprotokoll bildet die Schnittstelle zur Berechnung der Bewegungsdaten durch ein dafür konzipiertes Testprogramm des DLR. Die enthaltenden Daten liegen in Form von Binärdateien vor. Es werden alle Informationen über genutzte Kameraeigenschaften, Bewegungsschätzung und die Ergebnismengen der Korrespondenzanalyse bereit gestellt. Inhalt und Anordnung sind über ein entsprechendes Protokollformat definiert.

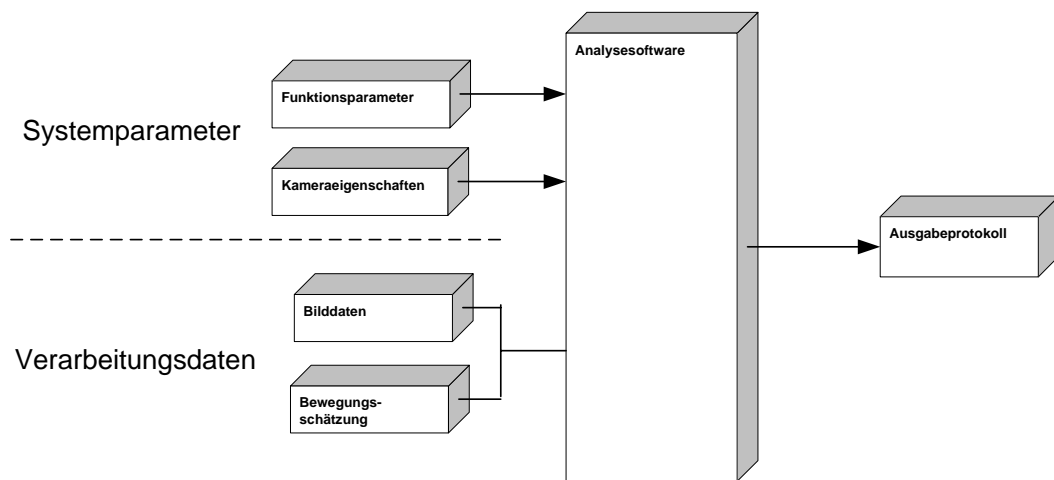


Abbildung 23: Kontext der Testsoftware

## 5.2 Erweiterung in der Klassenhierarchie

### 5.2.1 Die Kamera-Klasse

Über die Kamera-Klasse werden die intrinsischen Eigenschaften definiert. Ein Kameraobjekt enthält die Daten für die Kameramatrix und ihre Verzeichnisparameter. Neben Funktionen zum Zugriff auf die Membervariablen, werden die Ver- und Entzeichnung von Pixeln und die Skalierung von Kameraeigenschaften für die Bildpyramide implementiert. Die Funktion *distort* rechnet einen unverzeichneten Bildpunkt in die Koordinaten eines verzeichneten Bildpunkts um. Die Funktion *undistort* realisiert den umgekehrten Weg. Die Funktion *scaleMatrix* skaliert die Kameramatrizen, aufgrund des angegebenen Skalierungsfaktors.

CCamera
-A : CMatrix3 (Bibliothek) -k1 : double -k2 : double -k3 : double
+CCamera(void)() +~CCamera(void)() +setCam(CMatrix3<double> , double, double, double)() : void +setMatrix( CMatrix3<double> )() : void +setk1(double)() : void +setk2(double)() : void +setk3(double)() : void +getMatrixValue (void) () : CMatrix3 (Bibliothek) +getKParameter(unsigned int nr)() : double +distort(const double, const double, double &, double &)() : bool +undistort(const double, const double, double &, double &)() : bool +scaleMatrix( const float)() : void

Abbildung 24: Implementierung der Klasse Kamera

### 5.2.2 Die Klasse Epipolargeometrie

Die Klasse Epipolargeometrie beinhaltet alle implementierten Funktionen und Daten zur Berechnung der Epipolargeometrie und geometrisch bedingter Suchräume. Die Klasse enthält die Daten zweier Kameraobjekte, die komplette extrinsische Systemorientierung dT und stellt zusätzlich Informationen über Bilddimension, Korrelationsfenster, Suchbreite und die Suchbereichseingrenzung im Objektraum zur Verfügung.

Tabelle 1: Parameterbeschreibung der Klasse Epipolargeometrie

Variable	Beschreibung	Maximale Testwertebereiche
minDistance	Mindestabstand von Objektpunkten in Meter	$0,1 < \text{minDistance} < 1$
maxDistance	Maximalabstand von Objektpunkten in Meter	$\text{minDistance} < \text{maxDistance} < 1000$
dimX	Horizontale Bildgröße	Bildabhängig - typische Werte: 128,256,512,724,1024
dimY	Vertikale Bildgröße	Bildabhängig - typische Werte: 128,256,512,724,1024
corrwidth	relative Pixellänge des quadratischen Korrelationsfensters vom Bildpunkt aus: Anzahl der Pixel ergibt sich aus $(x*x)$ für $x=(2*corrwidth+1)$	$5 < \text{corrwidth} < 10$ für Korrelationsfenstergrößen von $121 < 441$ Pixeln
searchwidth	Relative Pixelangabe zur Verbreiterung des Suchbereichs. Linienbreite ergibt sich aus $(2*searchwidth+1)$	Je nach Suchrichtung: $0 < \text{searchwidth} < 5$ $0 < \text{searchwidth} < 5$

Die Bestimmung der Epipolarlinie aufgrund eines eingeschränkten Tiefenbereichs erfolgt durch die Funktion *calculateEpipolarline* über drei Parameterangaben. Diese berechnet die Geradengleichung und die Start und Endwerte der Epipolarlinie. Die Parameter der Geradengleichung werden als Membervariable zur internen Weiterverwendung und die Ergebnisse für Start- und Endwerte über Referenzen zurück gegeben. Als Übergabeparameter werden die Koordinaten des Ursprungspixels benötigt. Die Berechnung mit einem minimalen Abstand von 0 führt zur Berechnung der gesamten Epipolarlinie.

Zur Berechnung der gesamten Epipolarlinie stehen zwei weitere Funktionen zur Verfügung. Die Funktion *calculateFundamentalMatrix* berechnet die Fundamentalmatrix aufgrund der extrinsischen und intrinsischen Orientierung zweier Kameras. Unter Angabe des Quellpixels wird die Geradengleichung der Epipolarlinie auf Basis der Fundamentalmatrix durch die abgewandelte Funktion *calculateEpipolarlinewithF* berechnet. Start- und Endwert werden dabei aufgrund der Bilddimensionen bestimmt.

Mit Hilfe der Funktion *calculateLimitsforEpipolarline* werden die Koordinaten von Start und Endwerten einer Linie auf den aktuellen Bildbereich eingeschränkt. Bei verzeichneten Bildern entfällt jedoch die Ausführung der Funktion, da herangezogene Werte außerhalb der entzeichneten Bilddimensionen liegen können.

Die Berechnung aller Pixel der Suchraumlinie erfolgt durch die Funktion *calculateSearchList*. Die Variante mit drei Parametern berechnet aus Start- und Endkoordinaten alle auf der Linie enthaltenden Pixel aufgrund des Bresenham-Algorithmus. Durch Angabe einer Suchraumverbreitung (*searchwidth*) können auch parallel liegende Linien in den Suchbereich integriert werden. Die Ergebniserückgabe erfolgt in Form einer Liste mit allen Pixelkoordinaten des Suchraums.

Die Bestimmung der Suchräume für die Refining-Fenster des Pyramidenalgorithmus erfolgt mit der Funktion *calculateSearchList* durch vier Parameterübergaben. Übergabewerte ergeben sich aus dem Quellpixel und aus relativen Koordinatenangaben zur Erstellung des umliegenden Suchraums. Aus den Koordinatenangaben zur linken oberen Ecke und zur rechten unteren Ecke des Refining Fensters können alle umliegenden Koordinaten berechnet werden. Das Ergebnis wird ebenfalls als Pixelliste über eine Referenz zurückgegeben. Die Funktion kann dabei gegebenenfalls auch zur Bestimmung von Suchräumen über Fensterangaben, wie im Ursprungsprogramm wiederverwendet werden. Dies ist jedoch im Moment nicht implementiert.

Mit *saveEpipolarimage* steht eine Funktion zur Visualisierung des Ergebnisses zur Verfügung. Unter Angabe eines Zielverzeichnisses kann das Ergebnis zwischen Quellpixel und Suchliste in die betroffenen Bilder eingezeichnet und abgespeichert werden.

CEpipolargeometry
- K1 : CMatrix3<double> - K2 : CMatrix3<double> - deltaT : CMatrix4<double> - minDistance : double - maxDistance : double - DimX : int - DimY : int - corrwidth : int - Fundamentalmatrix : CMatrix3<double> - Epipolarline : sStraight - searchwidth : int
+CEpipolargeometry(void()) +CEpipolargeometry(double, double, mBitmapInfo, int, int, double, CMatrix4(dT), CCamera, CCamera()) +calculateFundamentalmatrix() : int +calculateEpipolarline(mPoint) : sStraight() +calculateEpipolarline(const mPoint, SubPixel &, SubPixel &) : int +calculateLimitsForEpipolarline(SubPixel &, SubPixel &) : int +calculateSearchList(SubPixel &, SubPixel &, vector <mPoint> &) : int +calculateSearchList(mPoint, mPoint, mPoint, vector <mPoint> &) : int +saveEpipolarimage(CPic *, CPic *, mPoint, line, string) : void +~CEpipolargeometry(void())

Abbildung 25: Implementierung der Klasse Epipolargeometry



### 5.3 Anpassungen in der algorithmischen Struktur

Die Steuerung des Gesamtalgorithmus wird in der Klasse `PointTracker` durch die Funktion *detectFeatures* umgesetzt. Dieser Programmteil wurde aufgrund ungünstiger Programmstrukturen und spärlicher Programmbeschreibung neu implementiert, wobei der Gesamtalgorithmus aus 3.2.7 weitestgehend erhalten bleibt. Änderungen in Datenhaltung und an Datenschnittstellen machen Neustrukturierungen des Programmcodes unvermeidlich.

Wichtige funktionale Änderungen ergeben sich aus der Umsetzung der Programmschnittstellen und den erweiterten Ausgaben. Die Initialisierung aller Parametereingaben erfolgt komplett aus der überarbeiteten Parametertextdatei. Die Daten werden zu Beginn des Programms ausgelesen und es folgt die Initialisierung aller allgemeingültigen Daten der gesamten Software. In der Parameterdatei werden 25 Parameter berücksichtigt, die vollständig zu Programmbeginn angegeben sein müssen. Eine Beschreibung der Parameterdefinitionen findet sich im Anhang.

Die Speicherung von Ergebnisdaten erfolgt wie gehabt in Form eines Ergebnisprotokolls für jeden Durchlauf und einer Textdatei zum Abschluss einer Testreihe. Ein einzelnes Ergebnisprotokoll beinhaltet wichtige Parameter und die Punktmengen eines Matching-Durchgangs über vier Bilder. Diese Ausgabedaten dienen zur späteren qualitativen Einschätzung der gefundenen Features. Ein vorhandenes Testprogramm des DLR ermöglicht es, durch die Ermittlung von Fehlmatches sowie durch die Rückrechnung der Bewegung aus den Punktmengen Aussagen über die Fehlerwahrscheinlichkeit zu gewinnen. Die zusätzliche Textdatei beinhaltet detaillierte Informationen über den Programmverlauf, wie Durchlaufzeiten, gefundene und gematchte Feature, sowie genutzte Parameter.

Debug-Funktionen zur Speicherung der gematchten Bilder werden komplett übernommen und wiederverwendet. Diese ermöglichen bei der Programmausführung in einem speziellen `DEBUG`-Modus, die untersuchten Bilder inklusive ihrer Matching-Ergebnisse in das Ergebnisverzeichnis abzuspeichern.

Die durch die Feature-Klasse implementierte Steuerung für die Merkmalsextraktion und das Matching, wird aufgrund der Optimierungsansätze, in der Funktionalität erweitert bzw. geändert. Die Durchlaufalgorithmen der Funktionen *featureExtractor* zur Merkmalsextraktion und *featureMatcher* für die Korrespondenzanalysen werden durch die Konzepte aus Kapitel 4 erweitert.

Zusätzlich werden Schnittstellen zum Einlesen der Bewegungsdaten und der Kamerakalibration geschaffen. Die Zwischenspeicherung der Kameramatrizen  $K_1$  und  $K_2$  samt Verzeichnisparameter und der extrinsischen Transformationen  $T_{tcp}$  und  $T_{CAM}$  als Memberfunktionen der Feature-Klasse erfolgt durch die Funktion *initCameraCalibration*. Zugehörig zu den Zeitpunkten zweier Stereoaufnahmen, werden aus einer Textdatei die relativen Orientierungsinformationen ausgelesen. Die gewonnenen Daten dienen zur Berechnung der relativen Systemorientierungen  $dT$  für das Intra- und Inter-Matching.

## **5.4 Funktionalitäten**

Die Software ermöglicht die Unterscheidung mehrerer Testfälle durch unabhängige Nutzung der Optimierungsansätze. Durch verschiedene Parametrierungen können die Eigenschaften und Auswirkungen nacheinander untersucht werden. Die Bestimmung von Suchräumen bei Korrespondenzanalysen erfolgt durch Berechnung der gesamten oder durch die distanzbeschränkte Epipolarlinie. Auswirkungen durch Ungenauigkeiten der Eingangsdaten können durch Verbreiterung des Suchraums berücksichtigt und untersucht werden. Frei wählbare Änderungen an der Parametrierung ermöglichen zusätzlich die Erweiterung des Algorithmus durch die Auflösungspyramide oder der Wiederverwendung von Merkmalen. Es besteht die Auswahl den vierten Matching-Schritt als Intra- oder Inter-Matching zu untersuchen. Die Nutzung von frei wählbaren Suchfenstern, wie im ersten Ansatz des Testprogramms, ist in dieser Version nicht mehr möglich. Vergleichsdaten können jedoch durch Nutzung der ersten Version erstellt und somit mit den aktuellen Ergebnissen validiert werden.

## **6 Testanalysen und Verifikation**

Mit Hilfe der modifizierten Testsoftware sollen verschiedene Versuchreihen helfen, das Verhalten der Optimierungsansätze zu klassifizieren. Schwerpunkte der Untersuchungen sollen Aussagen über die Auswirkungen auf das zeitliche Systemverhalten und die Qualität der Matching-Ergebnisse bilden. Ergebnisse aufgrund des alten Ansatzes können zur Verifikation herangezogen werden. Aufgrund möglichst gleicher Parametrierungen für Extraktions- und Matching-Algorithmen sollen zeitliche Einsparungen und Reduzierung des Arbeitsaufwands nachgewiesen werden. Zusätzlich soll die Verarbeitung verzeichneter Bildern aufgezeigt werden.

### **6.1 Versuchsbeschreibung**

#### **6.1.1 Testsoftware**

Die modifizierte Testsoftware „Testfeaturdetector.exe“ Version 2.0 ermöglicht es, Korrespondenzanalysen über aufgenommene Bildreihen von Stereokamerasystemen über Ansätze der Epipolargeometrie durchzuführen. Kamera- und Systemeigenschaften, sowie externe Bewegungsmessungen durch Angabe der Translation und des Drehwinkels, werden bei der Analyse berücksichtigt und fließen bei der Berechnung von Ergebnissen mit ein. Die Software ermöglicht sowohl Tests mit verzeichneten als auch mit unverzeichneten Bildreihen. In der aktuellen Version stehen die Merkmalsextraktoren KLT und Moravec in Kombination mit den Matching-Algorithmen NCC, SAD und SSD zur Verfügung. Zusätzliche Funktionalitäten sind mit dem Pyramidenansatz und der Wiederverwendung von Merkmalen implementiert. Je nach Anwendung kann beim vierten Matching-Schritt einer Korrespondenzanalyse zwischen Intra- und Inter-Matching gewählt werden.

#### **6.1.2 Testsystem**

Die Durchführung der Tests mit Hilfe der modifizierten Testsoftware „Testfeaturdetector.exe“ erfolgt auf einem handelsüblichen PC. Dessen Rechenleistung ist definiert durch einen AMD Athlon 64 X2 Dual Core Prozessor 3800+ mit 2,01 Ghz Taktfrequenz und 896 MB Arbeitsspeicher. Als Betriebssystem ist die 32-Bit Version von Windows XP Professional mit Service Pack 2 der Firma Microsoft installiert. Zur Speicherung der Ergebnisdaten steht eine 120 GB Festplattenpartition zur Verfügung.

### 6.1.3 Testreihen

Für Tests werden drei Bildreihen von Stereokamerasystemen aus Bewegungsfahrten in Innenräumen herangezogen. Zusätzlich zu den Bildszenen sind jeweils die Kalibrationsdatei und die Datei einer Bewegungsschätzung beigelegt. Die Testreihen unterscheiden sich grundsätzlich in zwei synthetisch erstellte Bildfolgen und einer realen Messreihe.

Die synthetischen Bildreihen basieren auf einer komplexen Simulation. Ausgangspunkt für die Simulation bildet ein virtuelles, texturiertes Raummodell. Durch diesen Raum kann eine Kamerafahrt unter Angabe der Bewegungsinformationen und der Kameraeigenschaften simuliert werden. Zur Erzeugung der Bilder wird ein virtuelles Lochkameramodell definiert. Die Eigenschaften ergeben sich aus Pixelgröße, Brennweite und Bildgröße. Die Bildaufnahme selbst kann zusätzlich durch Zeitpunkt und Blickrichtung beschrieben werden. Durch Zusammenspiel von Bewegung, Aufnahme Frequenz und Blickrichtung der Kamera kann somit eine Bildfolge innerhalb des virtuellen Raums erstellt werden.

Beide synthetischen Messreihen bestehen aus jeweils 80 Bildern. Die Aufnahmen haben eine Bilddimension von 724x724 Pixel und wurden mit einer Bildrate von 10 Hz erstellt. Die Bildsimulation I zeigt Aufnahmen eines achsparallelen Stereokamerasystems. Bei der Bildsimulation II sind die Kameras um jeweils fünf Grad zu einander verdreht. Alle Bilder liegen ohne Verzeichnungen vor. Die Kameraeigenschaften beider Kameras des Stereosystems wurden gleichgesetzt.



Abbildung 26: Synthetische Bildszene aus der achsparallelen Versuchsreihe

Tabelle 2: Kameraeigenschaften der synthetischen Bildreihen

Kameraeigenschaften	Bildsimulation I/II
Pixelgröße	10,6 $\mu\text{m}$
Brennweite	4,8 mm
Bildgröße	724x724

Die reale Messreihe wurde mit einem Versuchsaufbau aus zwei baugleichen Stereokameras des Modells DALSA 1M28-SA erstellt. Die aufgenommenen Bilder stammen aus einer Kamerafahrt bei einer Versuchsmessung in einem Büroraum. Die Bildreihe umfasst 529 Bildpaare und zeigt Bildausschnitte mit einer Größe von 724x724 Pixel. Durch die realen Kameraeigenschaften weisen die Bilder einen Verzeichnungsfehler auf. Die Bewegungsdaten wurden synchron zur Kamerafahrt in einer gesonderten Bewegungsmessung aufgenommen. Diese sind durch die absolute Position und Lage zum Bezugssystem beschrieben und für jeden Aufnahmezeitpunkt liegt jeweils ein Datensatz vor. Die Kameraeigenschaften, sowie die relative Orientierungen zum TCP und zwischen den Kameras sind durch Kalibration ermittelt und bekannt. Alle Datensätze unterliegen anhand der realen Eigenschaften gewissen Toleranzen, so das hier mit einigen Ungenauigkeiten zu rechnen ist.



Abbildung 27: Synthetische Bildszene aus dem realen Messdatensatz

Tabelle 3: Kameraeigenschaften der realen Messreihe

Kameraeigenschaften	Kamera I	Kamera II
Pixelgröße	10,6 $\mu\text{m}$	10,6 $\mu\text{m}$
Brennweite	4,8 mm	4,8mm
Bildgröße	724x724	724x724
Brennweite in Pixeln	459.991767	458.96201
Verzeichnungsparameter (K1/K2/K3)	K1 = -0.216688 K2 = 0.0609684 K3 = -0.008251;	K1 = -0.21509 K2 = 0.059883 K3 = -0.008251;

#### 6.1.4 Durchführung und Auswertung

Nachweis und Einfluss der einzelnen Optimierungsansätze werden durch Testfälle in verschiedenen Stufen dargestellt und analysiert. In der ersten Stufe werden die grundlegenden Ansätze der Epipolargeometrie denen des ursprünglichen Box-Verfahren gegenübergestellt. Vergleichbare Datensätze, basierend auf optimierten Suchbereichen beim Box-Verfahren und aus Berechnungen mit kompletter Epipolarlinie, werden erstellt. Auswirkungen auf Qualität und Arbeitsaufwand beider Verfahren werden miteinander verglichen und analysiert. In einer zweiten Stufe soll der Einfluss der einzelnen Optimierungsansätze nachgewiesen und klassifiziert werden. Ergebnisse aus der Korrespondenzanalyse mit Hilfe der gesamten Epipolarlinie bilden die Basisreferenz, um Auswirkungen der Optimierungsansätze mit Blick auf das zeitliche Verhalten und die Qualität der Merkmale zu beobachten und einzuschätzen. Eine exemplarische Prüfung verschiedener Operatoren, soll Optimierungsansätze durch Nutzung unterschiedlicher Algorithmen aufdecken. Die dritte Stufe beschäftigt sich mit Auswirkungen des Gesamtansatzes auf reale Bildpaare. Dabei soll ein erster Eindruck erarbeitet werden, wie sich reale Eingangsdaten auf die Bestimmung der Suchräume und Qualität der Features auswirken. Untersuchungen zur Verbreiterung des Suchbereichs wegen Ungenauigkeiten in den Eingangsdaten, sollen dazu in der Analyse berücksichtigt werden.

## 6.2 Gegenüberstellung des Epipolaransatzes zum Box-Verfahren

Die auf der Epipolargeometrie basierenden Ansätze, werden dem, auf festen Suchräumen basierendem, Ansatz des Box-Verfahrens gegenübergestellt. Die Ergebnisse dienen hierbei zur Einschätzung des Stands. Zur Gegenüberstellung werden Datensätze, beruhend auf dem Extraktor KLT und dem Matching-Algorithmus NCC, erstellt. Aufgrund steigender Rechenzeiten bei erhöhter Bildkomplexität im Box-Verfahren erfolgt die Prüfung mit Hilfe der achsparallelen Bildreihe aus Bildsimulation I. Die allgemeinen Versuchparameter sind dabei wie folgt definiert.

Tabelle 4: Parametrierung – Versuch I

	Ohne Pyramide	Pyramidenstufe 3
Extraktionsmethode:	KLT	KLT
Extraktionsschwellwert:	150	150
Matching Methode:	NCC	NCC
Matchingschwellwert:	0,95	0,9
Korrelationsfenster	17 x 17	17 x 17
Bildgröße	724 x 724	724 x 724
Bildanzahl:	80	80

Aufgrund der Suchfenstergrößen soll versucht werden den Arbeitsaufwand für jedes gefundene Feature annähernd zu bemessen. Die Suchbereiche für den Box-Ansatz ergeben sich aus den angegebenen Dimensionen der Suchfenster. Die Suchbereiche beim Epipolaransatz müssen durch zwei Fälle unterschieden werden. Im ersten Fall, dem Intra-Matching, sind die Suchräume aufgrund des achsparallelen Systems konstant, da alle Epipolarlinien parallel im Bild liegen. Eine Einschränkung von jeweils einem halben Korrelationsfenster, ergibt sich jedoch an den Außenbereichen. Beim Inter-Matching kann es jedoch zu unterschiedlichen Längen kommen, da Epipolarlinien, bedingt durch den Steigungswinkel, nicht durch das gesamte Bild verlaufen müssen. Aufgrund der durchgeführten Bewegung wird die Abweichung mit 10% angesetzt. Durch den Einsatz der AuflösungsPyramide in Stufe drei, können durch den Skalierungsfaktor von  $2^2$ , die Suchräume auf  $1/4$  reduziert werden. Ohne Nutzung des Pyramidenansatzes erfolgt die Berechnung aller genutzten Suchraumpixel aus  $2 \cdot \text{Suchraum}_{\text{intra}} + \text{Suchraum}_{\text{inter}}$ , wobei der Einfluss des Korrelationsfensters nicht berücksichtigt ist. Beim Pyramidenansatz muss zur Bestimmung des gesamten Suchraums ein sechsmaliger Durchlauf durch die Refining-Fenster in Betracht gezogen werden.

Tabelle 5: Suchfenstergrößen für Intra/Inter-Matching - Versuch I

	<b>Box-Ansatz</b>		<b>Epipolarlösung</b>	
	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>
<b><i>Suchfenstergrößen</i></b>				
Intra-Matching	86x11	22x3	708x1	177x1
Inter-Matching	141x141	35x35	708x1	177x1
Refining-Größe	-	5 x 5	-	5 x 5
<b><i>Hypothetischer Arbeitsaufwand bei Betrachtung eines Features</i></b>				
Suchbereichspixel	21773	1507	2124	504
Einsparung (Schätzung)	0	93%	90%	97%

Die Suchbereiche der Epipolarlösung müssen aufgrund ihrer Reduzierung auf eine Dimension nicht unbedingt kleiner sein als im Box-Verfahren. In Fällen mit minimaler Bewegung könnte im Box-Ansatz der Suchbereich kleiner ausfallen. Da entsprechende Bereiche aber nur willkürlich auf basierendes Vorwissen geschätzt werden können, ist das Verfahren in den meisten Fällen nicht praktisch einsetzbar. Die bisher theoretisch gemachten Angaben für Suchbereiche können lediglich auf ein einzelnes Feature bezogen werden. Der reale Arbeitsaufwand bei Matching Durchläufen hängt von der Anzahl der gefundenen und gematchten Features ab und ist durch die zufallsbedingten Bildstrukturen nicht deterministisch.

Tabelle 6: Matching-Ergebnisse zwischen Box-Ansatz und Epipolarlösung

	<b>Box-Ansatz</b>		<b>Epipolarlösung</b>	
	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>
<b><i>Feature-Anzahl</i></b>				
Extraktion (init)	27397	25222	28136	29181
1. Intra-Matching	24300	16170	23597	19165
Inter-Matching	22128	9699	19023	9739
2. Intra-Matching	21778	9254	18526	9260
Ergebnismenge	21778	8080	18526	7751
Fehlmatches	33	15	54	6
Fehler in %	0,15%	0,18%	0,29%	0,08%
Durchlaufzeit /s	5862.73 s	260,08 s	609.01 s	125,7 s



**Auswertung:** Die Ergebnisse aus der Gegenüberstellung zeigen, dass der Ansatz zur Dynamisierung der Suchbereiche aufgrund der Epipolargeometrie, sich in erster Linie positiv auf den Arbeitsaufwand auswirken. Die Durchschnittszeit zur Erfassung eines Features kann hierbei zum Vergleich herangezogen werden. Durch Normierung auf den Wert des Box-Verfahrens ohne Pyramide, ergeben sich relativ genau die Einsparungsverhältnisse, wie oben angenommen. Dieses Ergebnis ist jedoch kein Garant für qualitative Verbesserungen. So ergibt sich im Fall ohne Pyramidenansatz sogar eine Verschlechterung des prozentualen Fehlers durch erhöhte Fehlmatche. Ursache dafür bilden zufallsbedingte wiederkehrende oder merkmalschwache Bildstrukturen auf die der Matching-Algorithmus bei der Korrespondenzanalyse trifft. Dieses Phänomen wird durch Ausweitung des eindimensionalen Suchbereichs auf den gesamten Bildbereich, begünstigt. Features können bei ungünstiger Lage, nicht mehr eindeutig identifiziert werden und es kommt häufiger zu Fehlinterpretationen. Eindeutiger ist das Ergebnis im Pyramidenansatz. Durch die, bei der Skalierung, gröber werdenden Bildstrukturen werden merkmalschwächere Bildstrukturen von vorne herein nicht berücksichtigt. Dieser Umstand führt gleichzeitig zu einer geringeren Merkmalsmenge im Ergebnis. Die Ergebnismenge beruht nur noch auf den stärksten Features und lässt weniger Spielraum für Fehlmatche. Dies führt zur Erhöhung des Qualitätsmerkmals.

### 6.3 Einfluss der Optimierungsansätze

Durch sequentielles Hinzuschalten der einzelnen Optimierungsansätze, soll deren Einfluss und Effizienz untersucht werden. Als Testdatensatz wird die synthetische Bildreihe aus der Bildsimulation II herangezogen. Die Bildreihe besteht aus Aufnahmen mit jeweils um fünf Grad zueinander verdrehten Kameras. Somit ergeben sich jetzt auch beim Intra-Matching Epipolarlinien mit einer Steigung ungleich Null. Untersucht werden nacheinander die Einschränkung der Epipolarlinie, die Merkmalswiederverwendung und der Alternativansatz durch ein zweites Inter-Matching. Als Ausgangsreferenz dient ein Datensatz mit vollständiger Epipolarlinie.

Um Änderungen zwischen den einzelnen Ansätzen zu erfassen und somit vergleichbare Ergebnisse zu bekommen, werden auf alle Testfälle die gleichen Parameter angewendet. Im Vorfeld haben erste Testläufe ergeben, dass bei Nichtnutzung der Pyramide und einem Schwellwert von 0,95 bei der NCC, die Featureanzahl ausreichend konstant bleibt. Da mit einer Qualitätsbesserung in den Matching-Ergebnissen zu rechnen ist, wird der Schwellwert entsprechend angepasst. Für den Pyramidenansatz gilt dies jedoch nicht. Wegen der schwächer werdenden Bildstrukturen in skalierten Bildern, würden durch den stärkeren Schwellwert zu viele Merkmale herausfallen. Deshalb wird für den Pyramidenansatz der NCC Schwellwert mit 0,9 angesetzt.

Tabelle 7: Allgemeine Systemparameter

	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>
<b>Methoden</b>	<b>KLT/NCC</b>	<b>KLT/NCC</b>
Schwellwerte	150/0,95	150 / 0,9
Korrelationsfenster	17 x 17	17 x 17
Refining-Fenster	-	5 x 5
Bildanzahl:	80	80

Aufgrund unterschiedlicher Vorbedingungen bei Untersuchungen mit gesamter und eingeschränkter Epipolarlinie, müssen die Parameter für die minimale und maximale Distanz in beiden Fälle unterschiedlich angepasst werden. Im ersten Fall, der gesamten Epipolarlinie, gibt es praktisch keine Distanzangaben. Im Fall der eingeschränkten Epipolarlinie, sollen nur noch Objektpunkte mit einer Minstdistanz von 0,5 Meter und einer Maximaldistanz von 10 Metern herangezogen werden. Die Breite der Epipolarlinie wird mit einem Pixel angegeben. Der Suchbereich ist somit in beiden Fällen eindimensional.

Tabelle 8: Abweichende Parametrierungen für den Epipolarsuchbereich

	<b>Gesamte Epipolarlinie</b>	<b>Alle andere Fälle</b>
Minimale Distanz:	-	0,5
Maximale Distanz:	-	10
Breite der Epipolarlinie (Intra)	1	1
Breite der Epipolarlinie (Inter)	1	1

Da weder deterministische Aussagen über das Programmzeitverhalten noch bezüglich der Arbeitseinsparung gemacht werden können, ist eine absolute Messung über die Qualität und Auswirkungen der Optimierungsansätze nicht möglich. Aus der Gemeinsamkeit, dass alle weiteren Testdurchläufe den gleichen Systemeigenschaften zugrunde liegen, werden relative Aussagen über eine Referenzmessung angestrebt. Die Ergebnisse des Durchlaufs mit komplett durchsuchter Epipolarlinie und ohne Pyramidenansatz werden dafür zum Vergleich herangezogen.

Zur Beurteilung der Einsparungsverhältnisse bei Suchräumen können zwei Möglichkeiten in Betracht gezogen werden. Zum einen könnte eine Aussage zum zeitlichen Verhalten getroffen werden. Diese Möglichkeit ist durch das Programmverhalten innerhalb des Betriebssystems nur bedingt nutzbar und führt zu ungenauen Ergebnissen. Zum anderen ist es aber möglich die durchgelaufenen Suchpixel zu zählen. Berechnet man das Verhältnis aus den Ergebnismengen der korrespondierenden Features, so ergibt sich der Normierungsfaktor. Die Berechnung der realen Einsparung erfolgt zwischen der normierten Suchpixelanzahl des aktuellen Datensatz und der Suchpixelanzahl des ersten Datensatz ohne Epipolarlinie in Prozent.

Die Qualität der Matching-Ergebnisse, kann durch die Anzahl falscher Korrespondenzen ermittelt werden. Bezogen auf die Ergebnismenge der Features kann der Fehler prozentual auf die gematchten Ergebnisse beschrieben werden. Weiterhin sollen Unterschiede zwischen den Featuremengen untersucht und aufgedeckt werden.

Tabelle 9: Ergebnisse: gesamte und eingeschränkte Epipolarlinie

	<b>1. gesamte Epipolarlinie</b>		<b>2. eingeschränkte Epipolarlinie</b>	
	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>
<b><i>Featureanzahl</i></b>				
reused Feature	0	0	0	0
Extraktion	28042	29153	28042	29153
Intra-Matching	14974	12126	14722	12017
Inter-Matching	13371	10444	12813	10008
Intra-Matching	11997	9427	11581	9052
<b>Ergebnismenge</b>	11997	8308	11581	7983
<b>Fehlmatches</b>	49	15	29	12
<b>Fehler in %</b>	0,4%	0,18%	0,25%	0,15%
<b><i>Ø- Zeiten in Sekunden</i></b>				
Extraktion	0,65	0,04	0,66	0,04
Intra-Matching	2,77	0,74	0,67	0,25
Inter-Matching	1,41	0,34	0,27	0,12
Intra-Matching	1,34	0,32	0,32	0,13
<b>Zeit / Durchlauf</b>	6,17	1,44	1,92	0,54
<b>Matchingfrequenz</b>	0,2 Hz	0,7 Hz	0,5 Hz	1,8 Hz
<b><i>Ø- Arbeitsaufwand pro Bild (gezählte Suchpixel ohne Korrelation)</i></b>				
Intra-Matching	242665	62521	55276	18865
Inter-Matching	123384	28721	21654	9585
Intra-Matching	117913	26778	26398	10396
<b>Arbeitsaufwand</b>	483962	118020	103328	38847
<b><i>Arbeitsersparnis bei Suchpixeln normiert auf ersten Datensatz</i></b>				
Normierungsfaktor	1	0,693	0,966	0,665
Ø-Arbeitsaufwand (normiert)	483962	17303	106964	58417
<b>Ersparnis (%)</b>	0	64,8%	77,89%	87,9%

Tabelle 10: Ergebnisse: Merkmalswiederverwendung und 2. Inter-Matching

	<b>Matching-Recycling</b>		<b>Zweites Inter-Matching</b>	
	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>	<i>Ohne Pyramide</i>	<i>Pyramidenstufe 3</i>
<b><i>Featureanzahl</i></b>				
reused Feature	11250	6645	11735	6795
Extraktion	17408	22719	16903	29388
Intra-Matching	5680	6121	5356	12892
Inter-Matching	15035	10693	15203	10838
Intra-/Inter-Matching	13750	9689	14354	9861
<b>Ergebnismenge</b>	13750	8703	14354	8901
<b>Fehlmatches</b>	39	16	47	18
<b>Fehler in %</b>	0,28%	0,18%	0,33%	0,18%
<b><i>Ø- Zeiten</i></b>				
Extraktion	0.64	0.04	0.64	0.04
Intra-Matching	0.44	0.24	0.43	0.23
Inter-Matching	0.3	0.13	0.30	0.13
Intra-/Inter-Matching	0.36	0.14	0.26	0,13
Zeit pro Durchlauf	1.74	0.55	1.63	0,53
<b>Matchingfrequenz</b>	0,6 Hz	1,8 Hz	0,6 Hz	1,8 Hz
<b><i>Ø- Arbeitsaufwand pro Bild (gezählte Suchpixel ohne Korrelation)</i></b>				
Intra-Matching	33399	18033	32383	18132
Inter-Matching	24867	10296	25074	10398
Intra-/Inter-Matching	31004	11195	21832	9619
<b>Arbeitsaufwand</b>	89270	39526	79289	38149
<b><i>Arbeitsersparnis bei Suchpixeln normiert auf ersten Datensatz</i></b>				
Normierungsfaktor	1,146	0,725	1,196	0,741
Ø-Arbeitsaufwand (normiert)	77697	54519	66295	51483
<b>Ersparnis (%)</b>	83,9%	88,7%	86,3%	89,4%

**Auswertung:** Alle Optimierungsansätze zeigen positive Auswirkungen auf das Programmverhalten. So kann zur Bestimmung von Korrespondenzen durch Einschränkung der Epipolarlinie mit ca. 78% die größte Senkung des Arbeitsaufwands erreicht werden. Ähnlich große Auswirkungen zeigen sich auf die Programmpformance.

Nur liegt die prozentuale Zeitersparnis aufgrund der Messung ca. 10 % unter der realen Einsparung. Wie oben schon beschrieben ist dieses Ergebnis auf das nicht deterministische Zeitverhalten und die zusätzlichen Rechenkosten des Programms zurückzuführen. Der Einfluss der Programmperformance scheint auf diese Punkt zu konvergieren. So können durch die Merkmalswiederverwendung noch 5% und durch ein zweites Inter-Matching noch weitere 3% an Arbeitsaufwand eingespart werden.

Bei qualitativer Betrachtung der Ergebnisse fallen sofort entsprechende Einbußen im Zusammenhang mit den Ergebnismengen auf. Je größer die Ergebnismenge ist, desto größer wird auch der Fehler durch falsche Korrespondenzen. Einflüsse der Optimierungsansätze zur Verbesserung des Qualitätsverhaltens haben nur geringe Auswirkungen. So ist jedoch zu erkennen, dass sich nur die Optimierungsmaßnahme zur Einschränkung der Epipolarlinie positiv auf die Qualität auswirkt, wobei hier eine leichte Verringerung der Ergebnismenge zu verbuchen ist. Vergleicht man diese allerdings genauer mit der Ergebnismenge an Features bei der gesamten Epipolarlinie, so sind die Mengen nur geringfügig unterschiedlich.

Der Optimierungsansatz zur Merkmalswiederverwendung hat eine Steigerung der Ergebnismenge zur Folge. Nach einem Matching-Vorgang bleiben mehr Features bei geringerem Arbeitsaufwand gültig. Durch die Wiederverwendung kommt es allerdings zu Fehlerfortpflanzungen. Fehlmatche die nicht erkannt werden, bleiben gerade bei geringeren Bewegungen über mehrere Bildfolgen erhalten und nehmen so größeren Einfluss auf den gesamten Vorgang. Es ist daher anzunehmen, dass es bei einigen gezählten Fehlern, um einen solchen Fortpflanzungsfehler handelt. Dieser Umstand ist auch als Ursache für die Verschlechterung des qualitativen Ergebnisses des zweiten Inter-Matchings zu suchen. Durch Erhöhung der Ergebnismenge vergrößert sich auch der Fehler.

Werden im Gegenzug die Optimierungsansätze beim Pyramideneinsatz verglichen, ist ähnliches Verhalten zu beobachten. Durch die größeren Bildstrukturen bleibt jedoch eine kleinere Menge an gematchten Features übrig. Dabei wirkt sich der Fehler nur noch in geringerem Umfang aus. Ein Folgefehler nimmt aufgrund der höheren Eindeutigkeit der Features bei der Pyramide nur geringen Einfluss auf die anderen Optimierungsansätze. Zusätzlich ist jedoch auch hier eine Verbesserung der Qualität bei der eingeschränkten Epipolarlinie festzustellen.

Die Wiederverwendung von Features nimmt direkt Einfluss auf die Rechenkosten. Durch Wiederverwendung der Ergebnisliste sind entsprechende zeitliche Einsparun-

gen hauptsächlich beim Intra-Matching zu erkennen. Der Einfluss auf den Extraktionsoperator ist jedoch nur geringfügig. Dieses Ergebnis basiert auf den hohen Rechenkosten, die der KLT im Allgemeinen selbst verursacht. So ist es nicht möglich das zeitliche Verhalten dieses Operators bei konstanter Bilddimension dahingehend noch zu verbessern. Hier könnte sich jedoch der Einsatz anderer Operatoren mit geringerem Rechenaufwand als Vorteilhaft erweisen.

Das zweite Inter-Matching zeichnet sich durch die kleinen Bewegungen der Bildfolge als weiterer Optimierungsansatz aus. Bei der Nutzung des Pyramidenansatzes ergibt sich durch die kleinen Suchbereiche allerdings nur noch eine geringfügige Verbesserung, gegenüber der Pyramide bei der Merkmalswiederverwendung.

#### 6.4 Einfluss der Optimierungsansätze mit andere Operatoren

Eventuelle Einsparungsmöglichkeiten bei gleichbleibender Qualität können durch Nutzung von Operatoren mit geringerem Rechenaufwand erreicht werden. Hierzu werden exemplarisch Kombinationen aus KLT, Moravec und den Matching-Algorithmen NCC und SSD gegeneinander gehalten und verglichen. Ausgangslage bildet die Zusammenstellung des Gesamtalgorithmus unter Verwendung aller Optimierungsansätze. Es kommen die Distanzeinschränkung, die Merkmalswiederverwendung und das zweite Inter-Matching zum Einsatz. Die Festlegung der Parameter wird, wie folgt beschrieben, festgesetzt.

Tabelle 11: Parametrierung bei Gegenüberstellung der Operatoren

	KLT/SSD	MORAVEC/NCC	MORAVEC/SSD
Schwellwerte	150/0,02	150/0,95	150/0,02
Korrelationsfenster	17 x 17	17 x 17	17 x 17
Refining-Fenster	5 x 5	5 x 5	5 x 5
Bildanzahl:	80	80	80
Minimale Distanz:	0,5	0,5	0,5
Maximale Distanz:	10	10	10
Breite der Epipolarlinie (Intra)	1	1	1
Breite der Epipolarlinie (Inter)	1	1	1

Da unterschiedliche Operatoren auf die Größe der Suchbereiche keinen Einfluss nehmen, sollte bei allen Testfällen der Arbeitsaufwand relativ im gleichen Verhältnis zur Ergebnismenge stehen. Lediglich Auswirkungen bei der Weiternutzung von Fea-

tures sollten diesbezüglich näher betrachtet werden. Ergebnisse auf den Arbeitsaufwand sollen deshalb bezogen auf die Ergebnisse in 6.3 dargestellt werden. Zusätzlich wird ein Einblick auf das zeitliche Verhalten der einzelnen Operatoren im Bezug auf die Qualität der Features gegeben.

Tabelle 12: Matching-Ergebnisse bei verschiedenen Operatoren

	KLT/NCC	KLT/SSD	MORAVEC/NCC	MORAVEC/SSD
<b>Featureanzahl</b>				
reused Feature	11735	12895	14599	12544
Extraktion	16903	16515	27050	28073
Intra-Matching	5356	5474	2696	5179
Inter-Matching	15203	16238	15666	15490
Inter-Matching	14354	15286	14900	14582
<b>Fehlmatches</b>	47	36	85	34
<b>Fehler in %</b>	0,33%	0,24%	0,57%	<b>0,23%</b>
<b>Ø- Zeiten in s</b>				
Extraktion	0.64	0.64	0.02	0.02
Intra-Matching	0.44	0.36	0.38	0.34
Inter-Matching	0.3	0.27	0.31	0.27
Inter-Matching	0.36	0.24	0.27	0.23
<b>Zeit / Durchlauf</b>	1.74	1.51	0.97 s	0.85
<b>Matchingfrequenz</b>	0,6 Hz	0,7 Hz	1 Hz	1,1 Hz
<b>Ø- Arbeitsaufwand pro Bild (Suchpixel ohne Korrelation)</b>				
Intra-Matching	32383	31511	27570	29540
Inter-Matching	25074	26623	25136	25704
Inter-Matching	21832	23210	22253	22188
<b>Arbeitsaufwand</b>	79289	81344	74959	77432
<b>Arbeitsersparnis bei Suchpixeln normiert auf ersten Datensatz aus 6.2.2</b>				
Normierungsfaktor	1,19	1,27	1,24	1,21
Ø-Arbeitsaufwand (normiert)	66019	63849	60353	63730
<b>Ersparnis (%)</b>	86,3%	86,8%	<b>87,5%</b>	86,8%
<b>Zeitunterschiede normiert auf ersten Datensatz</b>				
Normierungsfaktor	1	1,06	1,03	1,01
<b>Zeit in s (normiert)</b>	1,74s	1,18s	0,78s	<b>0,7s</b>



**Auswertung:** Die Ergebnisse zeigen, dass die Auswahl der Operatoren elementar für die Qualität des Matchings ist. So ist festzustellen, dass durch Nutzung des SSD sich Qualität und zeitliches Durchlaufverhalten gegenüber dem NCC verbessern. Zusätzlich erhöht sich die Ergebnismenge, durch mehr wiedergefundene Features. Die Arbeitersparnis durch die Merkmalswiederverwendung kann dadurch geringfügig erhöht werden. Der Einsatz des Moravec-Operators bringt, durch seinen geringeren Rechenaufwand, bei der Extraktion eine enorme Zeitersparnis mit höherer Ausbeute gegenüber dem KLT. In Kombination mit dem NCC erhöht sich jedoch die Ergebnismenge kaum. Es ist festzustellen, dass die Anzahl der wiederverwendeten Feature steigt und somit diese Kombination die größte Aufwandseinsparung erbringt. Dies ist auf eine bessere Gleichverteilung der extrahierten Feature zurückzuführen. Betrachtet man jedoch den erheblich angestiegenen Fehler, so ist klar, dass die Qualität der Moravec Merkmale schlechter geworden ist. Die Zeitersparnis steht hierbei in keinem Verhältnis zur Qualität. Die Kombination Moravec/SSD kommt jedoch zu anderen Ergebnissen. Diese Kombination spiegelt das beste Zeitergebnis und die beste Qualität aller Operatorkombinationen wieder. Der SSD kompensiert dabei die Nachteile des Moravec, durch solideres Matching.

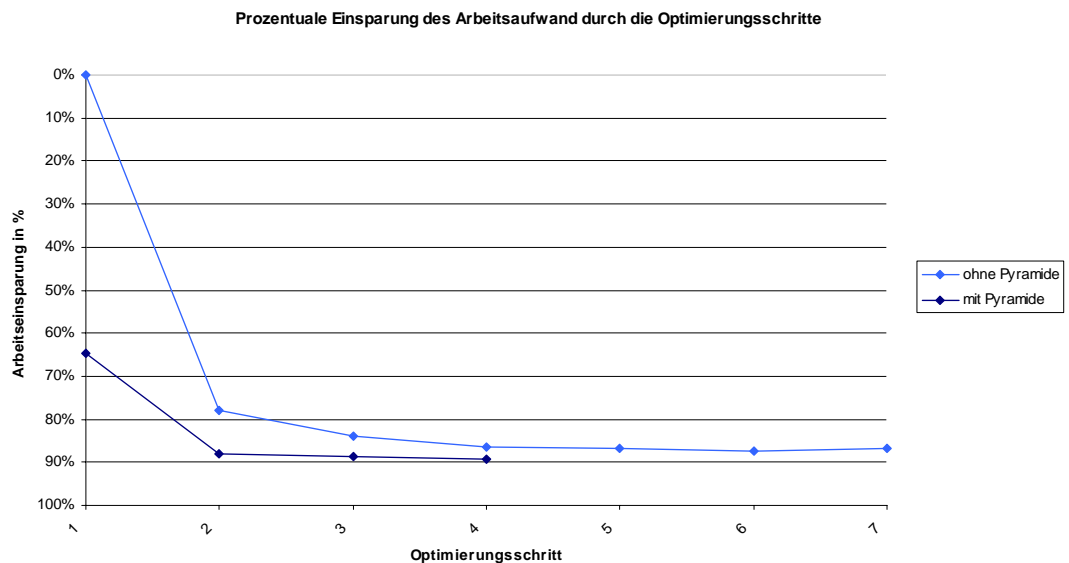


Abbildung 28: Arbeitseinsparung der untersuchten Optimierungsschritte in %

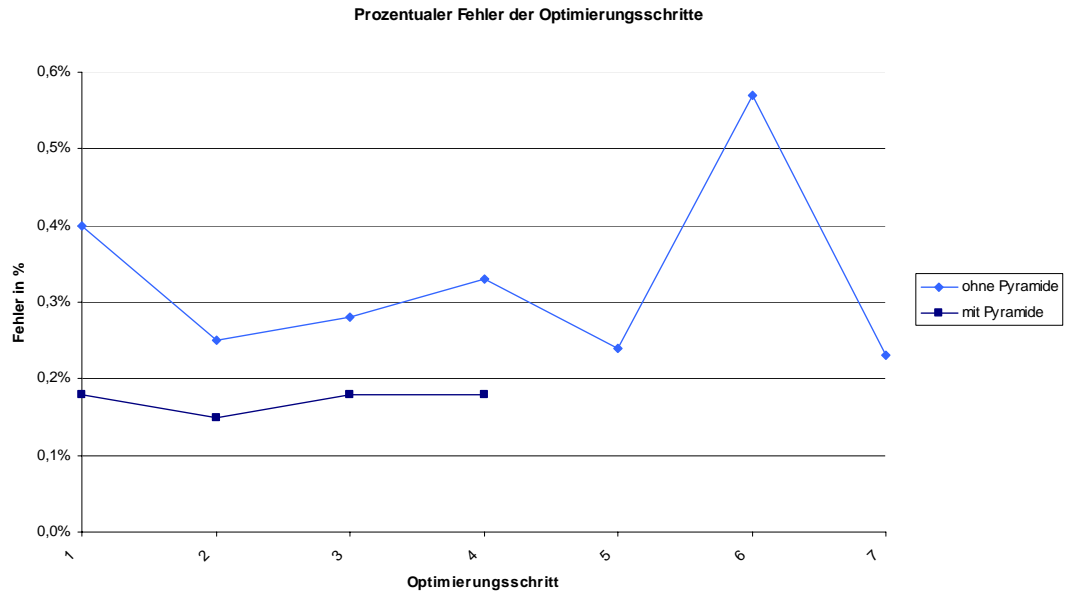


Abbildung 29: Fehlerverhalten der untersuchten Optimierungsschritte in %

*Legende Optimierungsschritte:*

- 1 gesamte Epipolarlinie (KLT/NCC)
- 2 eingeschränkte Epipolarlinie (KLT/NCC)
- 3 zusätzliche Nutzung der Merkmalswiederverwendung (KLT/NCC)
- 4 Erweiterung des Algorithmus durch zweites Inter-Matching (KLT/NCC)
- 5 Alle Optimierungsschritte (KLT/SSD)
- 6 Alle Optimierungsschritte (MORAVEC/NCC)
- 7 Alle Optimierungsschritte (MORAVEC/SSD)

## 6.5 Prüfung mit realen Bilddaten

Eine letzte Untersuchung soll das Verhalten der Optimierungsansätze auf die reale Bildfolge aufzeigen. Es wird versucht, durch optimale Parametrierung, Arbeitsaufwand und Qualitätsergebnisse der Korrespondenzen in Einklang zu bringen. Der Test erfolgt unter Einsatz aller Optimierungsansätze. Zur Erhöhung der Qualität für Features werden Schwellwerte für Extraktion und Korrelation entsprechend hoch angesetzt. Ziel ist es, bei noch ausreichend großer Featureanzahl eine hohe Qualität zu erzielen. Der Versuch wird mit den Operatoren Moravec und NCC durchgeführt. Der Moravec zeigt, wegen seines geringeren Rechenaufwands, bessere Durchlaufzeiten. Durch Erhöhung des Schwellwerts soll die Eindeutigkeit für Features nochmals verbessert werden. Für die Korrespondenzanalyse kommt, trotz besserer Ergebnisse des SSD, noch mal der NCC zum Einsatz. Aufgrund von Abweichungsfehlern bei der

Bewegungsmessung und durch Verzeichnungen muss die Epipolarsuchbreite erhöht werden. Der Pyramidenansatz kommt bei der Messung nicht zum Einsatz.

Tabelle 13: Parametrierung bei Untersuchung der realen Bildfolge

Methoden	MORAVEC/NCC
Schwellwerte	15/0,95
Korrelationsfenster	17 x 17
Bildanzahl:	317 (Startoffset: 210)
Minimale Distanz:	0,5
Maximale Distanz:	10
Breite der Epipolarlinie (Intra)	3
Breite der Epipolarlinie (Inter)	5

Tabelle 14: Matchingergebnisse bei realer Bildfolge

	MORAVEC/NCC
<b><i>Featureanzahl</i></b>	
reused Feature	41759
Extraktion	11828
Intra-Matching	45874
Inter-Matching	45273
Intra-Matching	45070
<b>Fehlmatches</b>	43
<b>Fehler in %</b>	0,1%
<b><i>Ø- Zeiten in s</i></b>	
Extraktion	0,02
Intra-Matching	0,22
Inter-Matching	0,05
Intra-Matching	0,04
Zeit pro Durchlauf	0.33
<b>Matchingfrequenz in Hz</b>	3 Hz
<b><i>Ø- Suchpixel (ohne Korrelation)</i></b>	
Intra-Matching	15421
Inter-Matching	3024
Intra-Matching	2293
Gesamtpixel	20738

**Auswertung:** Das Matching über die reale Bildreihe zeigt sehr gute Ergebnisse. Durch die Erhöhung des Schwellwerts beim Moravec Operator werden jetzt kleinere Featuremengen gematcht. Dieser Umstand führt zu einer erhöhten Durchlaufgeschwindigkeit, trotz verbreiteter Suchbereiche. Der qualitative Fehler in Form von Fehlmatches verringert sich ebenfalls. Mit 0,1% auf 45070 Features ist dies aus allen Versuchen das beste Ergebnis. Der Versuch zeigt, dass bei Erhöhung der Featureeindeutigkeit auch mit der Moravec/NCC Kombination gute Ergebnisse zu erzielen sind. Hierbei ist gut zu erkennen, wie die Qualität der Matchingergebnisse von der Qualität robusterer Merkmale der Extraktion profitiert. Durch den erhöhten Schwellwert besteht die extrahierte Menge aus eindeutigeren Features, die zu weniger Fehlmatches führen. Zusätzlich verringert sich die Merkmalsmenge die verarbeitet werden muss, so dass wiederum die Programmpformance profitiert. Aufgrund von Ungenauigkeiten in den Eingangsdaten z.B. durch Verzeichnungseigenschaften sind allerdings eindimensionale Suchräume nicht mehr ausreichend. Die somit vergrößerten Suchräume führen zu Erhöhungen beim Arbeitsaufwand, der sich durch die geringere Featureanzahl jedoch geringer auswirkt.

## **7 Zusammenfassung**

### **7.1 Erreichter Stand**

Die in dieser Masterarbeit vorgestellten Ansätze erlauben den optimierten Einsatz von Algorithmen zur Bestimmung von Korrespondenzen in Bildfolgen eines Stereosystems. Hierbei wirkt sich besonders die Einschränkung des Suchbereichs mit Hilfe der Epipolargeometrie positiv auf den Rechenaufwand aus. Die im Systementwurf beschriebenen Ansätze geben einen umfassenden Einblick über zu nutzende Schnittstellen und erforderliche Schritte zur Umsetzung der Suchbereichsbestimmung.

Aufgrund der Definition von Bezugssystemen können Bilddaten durch Euklidische Transformationen in Zusammenhang gebracht werden. Die Kenntnis über intrinsische und extrinsische Parameter erlaubt es mit Hilfe der Epipolargeometrie Abhängigkeiten zwischen zwei Kameraansichten zu formulieren. Durch die verloren gegangene Tiefeninformation ist jedoch eine punktgenaue Zuordnung von Bildkoordinaten des Quellsystems innerhalb des Zielsystems nicht möglich. Das Resultat stellt sich in Form der eindimensionalen Epipolarlinie dar, auf der alle möglichen Korrespondenzen abgebildet sind.

Durch Einschränkung des Tiefenbereichs für zu berücksichtigende Objektpunkte kann die Epipolarlinie auf einen Ausschnitt verkürzt werden. Somit werden nur noch Korrespondenzen abgebildet, die sich innerhalb dieser Grenzen befinden. Die Abbildung der Epipolarlinie auf das Pixelraster des Bildbereichs erfordert zusätzliche Korrekturmaßnahmen, die mit dem Bresenham-Algorithmus hinreichend ausgeglichen werden können.

Durch die Weiterverwendung bereits gefundener Merkmale und Korrespondenzen kann der Arbeitsaufwand bei der Merkmalsextraktion und beim ersten Intra-Matching zusätzlich verbessert werden. Hierzu muss die Ergebnismenge jedoch den Bildbereichen neu zugeordnet werden.

Bei der Verwendung von realen Bilddaten wird die Anwendung der linearen Ansätze der Epipolargeometrie durch eine Verzeichnungs Korrektur ermöglicht. Der verwendete Algorithmus bewirkt die Transformation von verzeichneten in unverzeichnete Bilddaten und umgekehrt. Aufgrund von Ungenauigkeiten bei den Eingangsdaten und durch den zusätzlichen Interpolationsschritt ist ein eindimensionaler Suchraum nicht mehr ausreichend. Mit Hilfe einer Suchraumverbreiterung kann dem Problem jedoch entgegen gewirkt werden.

Mit der Testsoftware können die Optimierungsansätze in ihrer Auswirkung und Qualität näher klassifiziert werden. Gegenüber dem Verfahren zur Bestimmung des eindimensionalen Suchraums durch die gesamte Epipolarlinie kann der Arbeitsaufwand durch zusätzliche Nutzung des eingeschränkten Tiefenbereichs verkleinert werden. Dieser Schritt hat neben Performancesteigerungen auch eine Erhöhung der Qualität zur Folge.

Die Merkmalswiederverwendung ermöglicht die Einsparung von Extraktions- und Matching-Schritte aufgrund schon bekannter Informationen. Hierbei werden Konstanz und Menge der Feature erhöht. Es ist jedoch festzustellen, dass entstehende Folgefehler innerhalb des Matchings sich negativ auf die Qualität auswirken können.

Durch die externen Bewegungsinformationen ist die Bestimmung von Suchbereichen auch beim Inter-Matching möglich. Ist die relative Position- und Lage bei der Bewegung kleiner als zwischen den Kameras ist es sinnvoll ein zweites Inter-Matching durchzuführen. Durch die kleiner werdenden Suchräume können weiteren Einsparungen beim Berechnungsaufwand gemacht werden.

Die Nutzung des Pyramidenansatzes ermöglicht eine höhere Qualität, bei geringerer Featureanzahl und kleineren Suchräumen. Durch die gröberen Bildstrukturen bei der Skalierung der Bilder, werden merkmalschwächere Features nicht berücksichtigt. Es kommt zu weniger Fehlmatches und somit zu einer generellen Erhöhung der Qualität.

Durch Gebrauch verschiedenener Operatorkombinationen sind zusätzliche Einsparungen durch einen geringeren Rechenaufwand möglich. So erweist sich die Moravec/SSD-Kombination mit den besten Zeit- und Qualitätseigenschaften aller untersuchten Operatoren. Hierbei ist festzustellen, dass gerade die Auswahl der Operatoren großen Einfluss auf die Qualität der Ergebnismenge hat. Die durch ungünstige Bildstrukturen hervorgerufenen Qualitätsfehler werden dabei zum größten Teil durch die Eigenschaften der Operatoren bestimmt.

## 7.2 Ausblick, Fortführung

Die Nutzung verschiedener Operatoren haben unterschiedliche Auswirkungen auf das Systemverhalten gezeigt. So müssen zur Erhöhung der Qualitätsergebnisse bei Korrespondenzanalysen weitere Untersuchungen zur Extraktion stabilerer Merkmalsmengen durchgeführt werden. Ziel dabei muss es sein, die zu untersuchende Featuremenge auf eindeutigere Merkmale zu stützen. Zusätzlich können durch Begrenzung ausreichend großer Extraktionsmengen die Durchlaufzeiten zusätzlich verringert werden. Ein möglicher Schritt kann dazu das Arbeiten mit stärkeren Schwellwerten bei der Extraktion und beim Matching sein.

Weitere Optimierungsschritte sind auch bei der Bestimmung von Suchraumgrößen möglich. So kann durch Kenntnis des korrespondierenden Bildpunkts einer zweiten Ansicht, die 3D-Information eines Objektpunkts vollständig zurück gewonnen werden. Dieser Ansatz ist aufgrund der vorhandenen Informationen aus beiden Bildern nach dem ersten Intra-Matching möglich. Eine weitere Suchraumbestimmung mit Hilfe der Epipolargeometrie wäre für entsprechende Punkte nicht mehr nötig. Durch die Tiefeninformation könnte die Bestimmung weiterer Korrespondenzen punktgenau aufgrund von kleinen Refining-Fenstern ermittelt werden

Durch die Ergebnisse dieser Arbeit können jetzt weitere Untersuchungen in Richtung Zustandsschätzung unternommen werden. Durch Rückrechnung der Bewegung aus der Ergebnismenge der Korrespondenzanalyse müssen Auswirkungen durch falsche Korrespondenzen im Zusammenhang zum Ergebnis jedoch näher klassifiziert werden. Ein weiterer Schritt in Richtung praktischer Umsetzung bildet die Entwicklung realer Schnittstellen zur Nutzung von Kamera- und Bewegungssensoren. Zusätzlich müssen Möglichkeiten und Grenzen in Hinblick auf die Echtzeitverarbeitung untersucht werden.

## Literaturverzeichnis

- [HaZ03] HARTLEY, RICHARD; ZISSERMAN, ANDREW:  
*Multiple View Geometry in Computer Vision, 2. Auflage*  
Cambridge University Press, 2003,ISBN: 0-521-54051-8
- [Jäh05] JÄHNE, BERND:  
*Digitale Bildverarbeitung, 6. Auflage*  
Springer Verlag, 2005,ISBN: 3-540-24999-0
- [Nif07] NISCHWITZ, ALFRED; FISCHER, MAX; HABERÄCKER, PETER :  
*Computergrafik und Bildverarbeitung, 2. Auflage*  
Vieweg, 2007,ISBN: 978-3-8348-0186-9
- [Rod04] RODEHORST, VOLKER:  
*Photogrammetrische 3D-Rekonstruktion*  
Wissenschaftlicher Verlag Berlin, 2004,ISBN: 3-936-84683-9
- [Ros06] ROSENTHAL, JEAN-CLAUDE:  
*Feature Detection und Matching Verfahren zur Position- und Lagebestimmung*  
Diplomarbeit (DLR), 2006
- [Schr05] SCHREER, OLIVER:  
*Stereoanalyse und Bildsynthese*  
Springer Verlag, 2005,ISBN: 3-540-23439-X
- [Zha98] Zhang, Zhengyou:  
*Determining the Epipolar Geometry and its Uncertainty: A Review*  
*International Journal of Computer Vision*, 27(2), 161– 198 (1998)
- [Zha00] Zhang, Zhengyou:  
*A flexible new technique for camera calibration*  
*IEEE Transactions on Pattern Analysis and Machine Intelligence*,  
.. 22(11):1330–1334, 2000.



## **Erklärung zur Urheberschaft**

Ich versichere hiermit, daß die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur bzw. Hilfsmittel ohne fremde Hilfe angefertigt habe.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Berlin, den 13.05.2008

(Eilke Santjer)

## **Anhang**

### **Anhangverzeichnis**

<b>Anlage A: Programmbeschreibung .....</b>	<b>75</b>
<b>Anlage B: Verschiedenes .....</b>	<b>79</b>
<b>Anlage C: CD-ROM.....</b>	<b>82</b>

## **Anlage A: Programmbeschreibung**

### **Programmkomponenten**

Bei der Testsoftware Testfeaturedetector.exe handelt es sich um ein Windows Konsoleprogramm. Es wurde in C++ mit Hilfe der Entwicklungsumgebung Visual Studio .NET 2003 der Firma Microsoft entwickelt und implementiert. Das Programm liegt in einer Release und in einer Debug Version dieser Arbeit bei. Bei der Release Version werden die Ausgabeprotokolle der Korrespondenzanalyse und ein Durchlaufprotokoll als Ausgabedateien erzeugt. Das Programm läuft dabei mit maximaler Performance. Die DEBUG Version ermöglicht zusätzlich die Speicherung der Testbilder mit eingezeichneten Matchingergebnissen. Dieser Umstand verschlechtert jedoch die Programmperformance durch zusätzlichen Overhead. Zur Ausführung des Programms werden generell folgende Komponenten benötigt:

- Testsoftware: Testfeaturedetector.exe (2.02)
- Parameterdatei (paramameter.txt)
- Bilddaten des Stereokamerasystems (Kamer 1+2 als bmp/jpeg)
- Datei mit zugehörigen Bewegungsdaten (t\_move.m)
- Kamerkalibration (cal\_stereo.m)

Bei der Ausführung des Programms ist zu beachten, dass die ausführbare Datei und die Parameterdatei in einem Verzeichnis liegen. Bei der Nutzung der Eingangsdaten (Bild- und Bewegungsdaten, Kalibrationsdatei) ist es sinnvoll diese in einer einheitlichen Verzeichnishierarchie anzuordnen. Die Ausgabe der Ergebnisdateien erfolgt bei Ablauf des Programms innerhalb des Bildoberverzeichnisses in einem Unterordner RESULTS.

### **Konfiguration und Parametrierung**

Bevor das Programm in Betrieb genommen werden kann, muss die Parameterdatei entsprechend der Pfade und Einstellungen angepasst werden.

### **Merkmalsextraktoren**

Die Funktionen der Merkmalsextraktoren KLT und Moravec wurden aus der ersten Version des Testprogramms integriert. Die Parametrierung zur Nutzung der Operatoren erfolgt durch Angabe der Methode, des Merkmalsgrenzwert und der Kachelgröße. Die Wiederverwendung von Features kann ein und ausgeschaltet werden. Gute Ausgangswerte zur Angabe der Merkmalsgrenze sind bei KLT >150 und für Moravec >12. Eine Erhöhung dieser Werte hat eine Verschärfung des Schwellwertkriteriums zu Folge.

## **Matching-Metriken**

Zur Korrespondenzanalyse stehen die Matching-Metriken NCC, SAD und SSD zur Verfügung. Die Konfiguration erfolgt durch Angabe des Matching-Grenzwerts, des Korrelationsfensters, sowie minimaler und maximaler Suchtiefe. Optional können Angabe zur Suchbreite unabhängig zum Intra- und Inter-Matching gemacht werden. Wertangaben für die Suchbreiten und das Korrelationsfenster, erfolgen immer relativ zum Ausgangspunkt. Somit ergibt sich z.B. eine relative Suchbreite von einem Pixel eine Gesamt-suchbreite von drei Pixeln. Durch Angabe der minimalen Distanz mit einem Nullwert, erfolgt die Berechnung des Suchraums über die gesamte Epipolarlinie. Die Einschränkung auf den Tiefenbereich kann somit abgeschaltet werden.

## **Kalibrations- und Bewegungsdaten**

Die Integration der Kamerakalibrierung und der Bewegungsdaten erfolgt entsprechend definierter Textdateien. Zur Nutzung der Dateien sind der Verzeichnispfad und der Dateiname anzugeben. In der Regel liegen die entsprechenden Dateien im Verzeichnis der Bilddaten

## **Einstellungen zur Pyramide**

Die Einstellungen zur Pyramide setzen sich aus dem Pyramidenlevel und der Größe des Refining-Fensters zusammen. Der Wert 1 für den Pyramidenlevel parametrisiert den Durchlauf ohne Pyramide. Die Größe des Refining Fensters wird als Differenzwert angegeben. So ergibt eine Größe von 2 ein Fenster von 5x5 ( $2 \cdot \text{Größe} + 1$ )

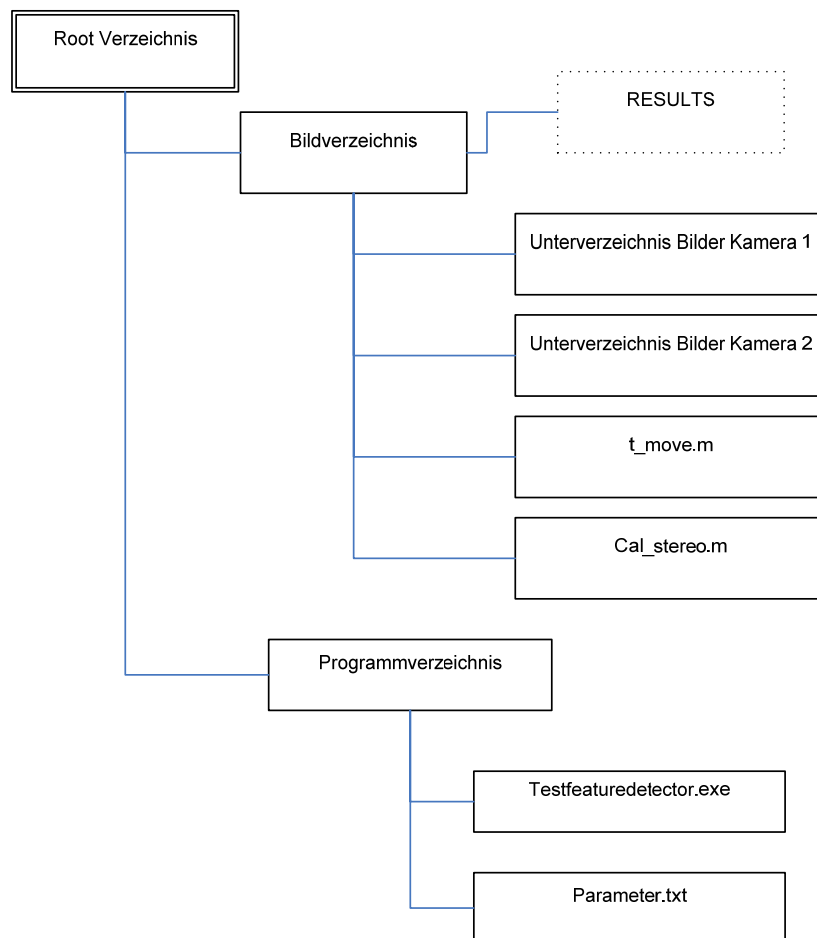
## **Nutzung der Bilddaten**

Bilddaten müssen aus einem Dateiprefix, der Bildnummer und dem Dateisuffix bestehen. Der Verzeichnispfad beschreibt den Ort der Dateien im Dateisystem. Bilddaten werden zwischen erster und zweiter Kamera unterschieden. Für beide gilt die Angabe des Dateinamens durch das Prefix. Die Angabe kann dabei mit einem Unterverzeichnis kombiniert angegeben werden. Der Startframe beschreibt die Bildnummer des ersten Bildes einer Bildreihe. Auch diese Information wird für die erste und zweite Kamera getrennt angegeben. Das Suffix gibt die Dateiendung und Art der Bilder an (bmp/jpeg). Die Angabe des Bildoffsets wird den Bildnummern zugerechnet. Dies ermöglicht den Start einer Untersuchung bei einer späteren Framenummer. Über die Filterangabe wird der Glättungsfilter angegeben. Es kann zwischen Binomial- und Medianfilter gewählt werden.

## Steuerparameter

Zur Steuerung des Programmalgorithmus stehen weitere Parameter zur Verfügung. So kann durch den Parameter `Feature_Recycling` die Merkmalswiederverwendung ein und ausgeschaltet werden. Mit Angabe des Parameters `Last_Matchingpath` wird gewählt, ob ein zweites Intra- oder Inter-Matching angewendet werden soll.

Beispiel für eine mögliche Verzeichnishierarchie des Testprogramms :



### Aufbau der Parameterdatei:

<b>0_FEATUREMETHOD</b> <KLT, MORAVEC>	//Extraktionsmethode
<b>1_EXTRACTION_THRESHOLD</b> < 150 oder größer>	//Schwellwert für die Extraktion
<b>2_NUMBER_OF_FEATURES</b> -1	// max. Features (nicht implementiert)
<b>3_GRID_SIZE</b> <z.B. 30 >	//Kachelgröße z.B. 30x30
<b>4_MATCHING_METRIC</b> < NCC, SSD, SAD >	//Matchingmethode
<b>5_MATCHING_CORRELATIONWINDOW</b> <z.B. 8 >	// halbe Korrelationsfenstergröße
<b>6_MATCHING_THRESHHOLD</b> < 0.1(NCC)/0,95(SSD)>	//Schwellwert Matching
<b>7_MINDISTANCE</b> <z.B. 0.5 >	//Min. Distanz Objekte
<b>8_MAXDISTANCE</b> <z.B. 10 >	//Max. Distanz Objekte
<b>9_INTRALINEWIDTH</b> < z.B.0,1,2...>	//Halbe Suchraumbreite Intra
<b>10_INTERLINEWIDTH</b> < z.B.0,1,2..>	//Halbe Suchraumbreite Inter
<b>11_CALIBRATIONFILE</b> < Verzeichnis und Dateiname>	//Kamerakalibration
<b>12_MOVEFILE</b> < Verzeichnis und Dateiname>	//Bewegungsdaten
<b>13_PYRAMIDE_LEVEL</b> <z.B. 1,2,3..>	//Pyramidenlevel
<b>14_REFINING_SIZE</b> <z.B. 2>	//halbe Größe Refining Fenster
<b>15_OFFSET_FROM_STARTING_POINT</b> <z.B.0>	//Bildoffset
<b>16_IMAGEDIRECTORY</b> < Verzeichnis >	//Bilderverzeichnis
<b>17_FileName_CAM1_PART1</b> <z.B. head_Cam_left/F >	//mögliches Subdirectory + Prefix Bilddatei
<b>18_FileName_CAM2_PART1</b> <z.B. head_Cam_right/F >	//mögliches Subdirectory + Prefix Bilddatei
<b>19_FileName_CAM1_PART2_STARTFRAME</b> <0,1..>	//Framennummer 1. Kamera
<b>20_FileName_CAM2_PART2_STARTFRAME</b> <0,1..>	//Framennummer 2. Kamera
<b>21_FileName_PART3_SUFFIX</b> <.bmp;.jpg>	//Dateisuffix
<b>22_FILTER_MEDIAN_BINOMIAL</b> <MEDIAN BINOMIAL>	//Glättungsfilter
<b>23_FEATURE_RECYCLING</b> <ON OFF>	//Merkmalswiederverwendung
<b>24_LAST_MATCHINGPATH</b> <INTRA INTER_RIGHT>	//Pfad für zweites Inter-Matching

# Anlage B: Verschiedenes

## 1 Klassenhierarchie des Versuchsprogramms



## 2 Kalibrationsdatei

data item	comment	type
%<comment with first item %>		Textstring
cam1.A = [a <sub>11</sub> a <sub>13</sub> a <sub>13</sub> a <sub>21</sub> a <sub>21</sub> a <sub>21</sub> 0 0 1];		9 x double
cam1.k1 = <k1(double)>;		double
cam1.k2 = <k2(double)>;		double
cam1.k3 = <k3(double)>;		double
cam2.A = [a <sub>11</sub> a <sub>13</sub> a <sub>13</sub> a <sub>21</sub> a <sub>21</sub> a <sub>21</sub> 0 0 1];		9 x double
cam2.k1 = <k1>;		double
cam2.k2 = <k2>;		double
cam2.k3 = <k3>;		double
Tcam = [a <sub>11</sub> a <sub>13</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> 0 0 0 1];		12x double
Ttcp = [a <sub>11</sub> a <sub>13</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> 0 0 0 1];		12x double

## 3 Bewegungsdatei

data item	comment	type
%<comment with first item %>		Textstring
Delta=[ time a <sub>11</sub> a <sub>12</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> a <sub>31</sub> a <sub>32</sub> a <sub>33</sub> a <sub>34</sub> 0 0 0 1	start of move file +	17 x double
time a <sub>11</sub> a <sub>12</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> a <sub>31</sub> a <sub>32</sub> a <sub>33</sub> a <sub>34</sub> 0 0 0 1	1. dataset	17 x double
....	2. dataset	17 x double
time a <sub>11</sub> a <sub>12</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> a <sub>31</sub> a <sub>32</sub> a <sub>33</sub> a <sub>34</sub> 0 0 0 1	...	17 x double
time a <sub>11</sub> a <sub>12</sub> a <sub>13</sub> a <sub>14</sub> a <sub>21</sub> a <sub>22</sub> a <sub>23</sub> a <sub>24</sub> a <sub>31</sub> a <sub>32</sub> a <sub>33</sub> a <sub>34</sub> 0 0 0 1 ]	n-1 dataset	17 x double
	N dataset+ End of	17 x double
	move file	



#### 4 Ausgabeprotokoll

Header length: 416 byte

Data length:  $n * 11 * 4$  byte

data item	units	type
sensor size (u,v)	Pixel	2 x double
pixel size	m	double
camera matrix I ( $a_{11}, a_{13}, a_{13}, a_{21}, \dots$ )		9 x double
distortion camera I ( $k_1, k_2, k_3$ )		3 x double
camera matrix II ( $a_{11}, a_{13}, a_{13}, a_{21}, \dots$ )		9 x double
distortion camera II ( $k_1, k_2, k_3$ )		3 x double
$T_{cam} (a_{11}, a_{13}, a_{13}, a_{14}, a_{21}, \dots)$		12x double
$T_{tcp} (a_{11}, a_{13}, a_{13}, a_{14}, a_{21}, \dots)$		12x double
Number of object points		double
object point 1, location1, camera 1 (u,v)	pixel	2 x float
object point 1, location1, camera 2 (u,v)	pixel	2 x float
object point 1, location1, correlation	-	float
object point 1, location2, camera 1 (u,v)	pixel	2 x float
object point 1, location2, correlation	-	float
object point 1, location2, camera 2 (u,v)	pixel	2 x float
object point 1, location2, correlation	-	float
...		
object point n, location1, camera 1 (u,v)	pixel	2 x float
object point n, location1, camera 2 (u,v)	pixel	2 x float
object point n, location1, correlation	-	float
object point n, location2, camera 1 (u,v)	pixel	2 x float
object point n, location2, correlation	-	float
object point n, location2, camera 2 (u,v)	pixel	2 x float
object point n, location2, correlation	-	float

## **Anlage C: CD-ROM**